

THIS WEEK

EDITORIALS

TOOLBOX *Nature's* latest guide to keep you up to date on technology **p.6**

WORLD VIEW Global power lists are unfair to Africa's scientists **p.7**



ROCKS Ice and GPS solve the mystery of shifting stones **p.8**

There is life after academia

With high numbers of postdocs emerging from universities, prospective PhD students must be prepared for the fact that they will probably not end up with a career in research.

When is a scientist not a scientist? To many people in the wider world, anyone with a PhD in nuclear physics or molecular genetics would probably still earn the moniker — whether they work as a researcher in an Ivy League laboratory or as a full-time television presenter. To those on the inside, the term is usually reserved for those with science as a professional vocation. Science is something that one ‘does’, not merely something that one has studied. To the hardliners, someone with postgraduate training who then leaves the lab to work elsewhere is judged to have lost their way, or been seduced by the dark side. It is no coincidence that employment opportunities for postgraduates outside research are often labelled ‘alternative’ careers.

Nature, of course, would not come out each week were it not for the scientists who publish here. But the journal's existence also relies on former researchers on the staff who are following alternative careers, whether they work as manuscript editors or journalists, or fill other essential roles. There is scarcely a page of this issue that was not in some way touched by an ex-scientist — please do not call us ‘failed’ — including many who did PhDs and postdocs, and even the odd former faculty member. This piece alone was brought to you with the help of three postgraduate qualifications.

We would say this, wouldn't we, but it would be a huge mistake to assume that those who jump from the research career ladder were unable to hack it.

For a glimpse of why three outstanding scientists left academia and how their training influenced them in their careers and lives, turn to page 20 for our News Feature ‘The ones who got away’. Although ‘data’ is not the plural of ‘anecdote’, their stories speak to the fact that plenty of the most capable postdocs will never become group leaders, and that many will never do research again once they finish their degree or PhD.

The picture is different across the world. Take motive: more students in Britain than in the United States and Germany are likely to begin PhDs without being committed to a career in research (if only because they will be able to put the title ‘Dr’ on their credit card in three years, rather than in six or seven).

REBRAND

Figures collected by the US National Science Foundation show that nearly 20% of Americans with science PhDs were not working in science in 2010, and full-time jobs in academia are proving harder to come by. A telling infographic (go.nature.com/vhlewm) from the American Society for Cell Biology points out that less than 10% of the 86,000 current biology PhD students in the United States will become tenure-track faculty members. ‘A faculty job is an ‘alternative’ career,’ the graphic cheekily states, while noting that more than half of PhD students rank a professorship as their most desired career.

You can look at this glut in a couple of ways. *Nature's* 2011 special, ‘The future of the PhD’, presented all these extra doctors as a problem in need of a fix (see nature.com/phdfuture). In one story, ‘The PhD

factory’, we noted that the world was producing more postdocs than ever, with the implicit suggestion that fewer of them would probably be a good thing: ‘Supply has outstripped demand and, although few PhD holders end up unemployed, it is not clear that spending years securing this high-level qualification is worth it for a job as, for example, a high-school teacher,’ we wrote (see *Nature* **472**, 276–279; 2011). Another piece offered suggestions for revamping the education of future scientists — from earlier independence to online courses to ditching the PhD entirely (see *Nature* **472**, 280–282; 2014).

“Universities should do more to help their PhD students to gain skills and contacts that will come in handy beyond academia.”

But instead of culling graduate students or abandoning the PhD, why not rebrand it? Rather than being a first rung on a ladder that ends with tenure-track professor (unless you tumble off), doctorates could be treated more like a trail that feeds through to a number of different paths (some easier, some harder, some even rather scary).

Many enlightened group leaders already see their roles as mentors in this way. When contacted about participating in ‘The ones who got away’, one group leader who did not end up in the story suggested that we switch to the headline ‘How scientists have had super careers in areas other than lab research’. Another scientist we contacted runs an annual alternative-careers workshop, so that the PhD students and postdocs in her lab can see that there is life after academia.

The burden should not fall only on group leaders and mentors. When universities are wooing prospective postdocs, they should make it clear that most will not end up as faculty members, while pointing out all the other careers that their graduate students have gone on to pursue.

Universities should also do more to help their PhD students to gain skills and contacts that will come in handy beyond academia, for instance by making it easier for students to take courses in other schools and departments. This is not to say that graduate school should turn into a trade school. Designing, conducting and interpreting experiments teaches young scientists how to do more than move minuscule amounts of water around and calculate *P*-values. Soroosh Shambayati, a chemist turned investment banker whom we profile in the News Feature, found that his talents for setting up multi-step chemical reactions prepared him for trades in derivatives.

Armed with this kind of information and support, would-be graduate students can make more-informed decisions about whether a PhD is right for them, whatever their goals. It would be misguided for new graduates to see their PhD purely as a career stepping stone, but it would be equally wrong to deny PhDs to all but the most committed academics.

Certainly, many of those who pay for postgraduate training are sanguine about where these ‘scientists’ end up working. As National Institutes of Health director Francis Collins told *Nature* in a recent interview: ‘They are not alternative careers, they are just careers.’ ■

Heavenly homes

The discovery of our Galaxy's place in the Universe adds detail to our address.

As Tim Radford, former science editor of *The Guardian* newspaper, noted in his 2011 book *The Address Book*, inquisitive schoolchildren the world over have a certain ritual. Handed a new exercise book and asked to add their name and address, they do so on a Universal scale. House number, street, town and country, postcode even, are followed by their designated continent and the name of our planet. Up the cosmic scale they go, noting the Solar System then the Milky Way, before offering the final identifier: the Universe.

It sounds precise; it is anything but. Most difficult for a deep-space postal service would be the first jump, from the infinite stretch of the known Universe down to our local Galaxy, the Milky Way. Well, things just got a little easier. (Although perhaps not for Radford, who may need to update his book.) This week, scientists add a new line to our planetary coordinates: the Laniakea galaxy supercluster.

Do not bother googling the name. It really is brand new, coined by an international group of astronomers on page 71 of this issue. Our place in the Universe, for so long one of the core mysteries of human existence that scientists and this journal are dedicated to unravelling, just got a little clearer. Laniakea, the scientists write, is our home supercluster, the one in which the Milky Way resides.

What kind of home is it? It is big — some 160 million parsecs across. Although not as big as some superclusters, it is the largest in our local neighbourhood, which is surprisingly crowded given the vast emptiness of most of the cosmological void. Laniakea has

several supercluster neighbours, including Coma, Perseus–Pisces and Shapley. (Together they make what? A super-supercluster? A hypercluster?)

It is a home that has been hiding in plain sight, colossal and all around us, yet unnoticed by previous astronomical surveys. As Elmo Tempel discusses in a News & Views article on page 41, this is probably because the boundaries of superclusters are tricky to pin down, even for astronomers.

“It is a home that has been hiding in plain sight, colossal and all around us, yet unnoticed by previous astronomical surveys.”

Laniakea was finally spotted with the help of what Tempel calls a “nifty algorithm” that helped the astronomers to turn incomplete measurements of the motion of galaxies into a map of the distribution and dynamics of cosmic matter. Their map shows galaxy superclusters as hotspots — basins of attractions in fields of velocity flow — that can be hived off from their surroundings.

It is a local map. The nifty algorithm is limited because it depends on direct measurements from Earth of how rapidly galaxies recede owing to cosmic expansion. The rest of the Universe — those galaxies far, far away — remains uncharted territory, for now.

Still, the Laniakea survey is more than cartographical and geographical information. It reveals details of the large-scale structures that surround the Milky Way (best viewed in a video available here: go.nature.com/hpjzwh), which should help astronomers to close in their determinations of cosmological parameters such as the density of dark energy, the hidden power believed to push the Universe away from us.

The name Laniakea has Hawaiian roots, and roughly translated means spacious heaven. It is a beautiful address to have. And one that comes just in time for the new school year and a new curious generation. ■

The digital toolbox

A new section of Nature examines the software and websites that make research easier.

Asked to list the essential tools of a scientist's trade, most would probably think first of hardware: the microscope, the telescope, the mass spectrometer, the genome sequencer, the test-tube. But just as important to today's data-wranglers are software — Excel, ChemDraw, MATLAB — and the programming languages used to create it, such as Python, R and SQL. Such tools are integral to modern research practice, whether for analysing or visualizing data, sharing files, collaborating, writing up papers, publishing, searching the literature or simply organizing one's work. And although software engineers have often overlooked science in favour of more lucrative markets — think Flappy Bird, Instagram and iTunes — software, websites and apps designed specifically for researchers are blossoming.

Partly in response to this flourishing sector, *Nature* this week introduces a new section to help readers keep up to date. The Toolbox pages will collect the journal's writing on software tools and websites that researchers use to work more efficiently, or in new ways. Find them online at nature.com/toolbox, and monthly in print.

Barely a week goes by without the appearance of a website offering improved research productivity, or the launch of a start-up firm hoping that its unique idea will change scientific workflow. Toolbox will aim to guide the perplexed through the maze of sites and programs, discussing their similarities and distinguishing points. But it will be

a community-driven resource, with scientists in various fields who work heavily with data or programs offering thoughts on their most commonly used software. The site will also collect *Nature's* writing on the broader context of online research — from open data to citizen science and crowd-funding.

In this issue, for instance, the section reviews a recent trend: the emergence of ‘recommendation engines’ that sift the flood of literature so that scientists can find relevant papers and information (see page 129). Future articles will include a look at the iPython interactive-computing project and its applications for scientists, and an examination of websites that promise to help researchers to collaborate on research papers.

A little software literacy can make any researcher's daily life more efficient. Version-control systems such as Git, for instance, help to record changes made to files, to allow recall and analysis of past and evolving work. Websites such as GitHub (a favourite of software engineers, but increasingly of scientists too) build on these platforms to help researchers to work collaboratively on a research paper, or to ensure that data analysis is clear and reproducible.

Such programming tools are already the daily bread of the data scientists, bioinformaticians and climate modellers among *Nature's* readers. But jargon can make them off-putting for non-coders. At the same time, it can be hard to tell which of the more polished, graphically friendly software packages are worth investing your time in. So on the Toolbox website, scientists will be able to share their recommendations for particular software (both commercial and free). As a

taster, this week the Software Carpentry movement, which teaches basic software skills to researchers, explains its motivations and operations. A bad workman may blame his tools; a good scientist needs to keep track of them. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunqv

IZETTE GREYLING



Judge research impact on a local scale

Metrics that give a global overview risk sidelining science in developing nations, argues Casparus J. Crous.

The latest global list of the world's most highly cited scientists was released in June and is likely to act as a focal point for promotion and funding decisions. Creators of such metrics — in this case the company Thomson Reuters — argue that they take no responsibility for how their data are used. But the reality here in South Africa, as in many other countries, is that they feed into official judgements of performance and discussions of how to allocate resources.

To me, an early-career African scientist, the latest list makes disturbing reading (highlycited.com). Searching across all entries, only 11 African scientists made it onto the list — just 0.3% of the total. Of these, only six have primary addresses in Africa. The list is dominated by the United States (more than half of all scientists featured), followed by a handful of European nations, as well as China and Japan. Most developing countries have similarly bleak showings, except Saudi Arabia, which has more scientists on the list — all from a single university — than any other nation except the United States and the United Kingdom.

If this is a true reflection of the world's leading minds, what does it mean for an aspiring scientist in Africa? It might be subjective, but if you have an interest in furthering your career, how could this easily accessible and well-published list influence you? It tells an enquiring young scholar little about current research growth or potential in Africa.

The value of role models for stimulating young people to follow a certain career is well known. When I reflect on my choice to venture into biology, I recall reading about the first human heart transplant, by Christiaan Barnard in South Africa in 1967. I was intrigued by this 'magic'. It wholly challenged my perception of life as I knew it; how on Earth was this possible? Today, considering the highly-cited list, I was confronted with a new question: is there any value in continuing an academic career in Africa when all the science of 'impact' is done elsewhere?

Arguably, there has never been a better time to be a budding scientist. With advances in technology, and the great increase and connectivity of available knowledge, I am part of a generation with unrivalled tools for investigating some of the world's most complex questions. Moreover, with various factors driving international collaboration, geographical and disciplinary boundaries no longer stand in the way of attacking evermore complicated questions. Fundamentally, young scientists should all strive to learn from the best researchers, in countries offering the highest scientific impact. The global scientific impact metric provides an attractive map with which to orient such an endeavour.

Excellence should be recognized and rewarded, especially when research funds are under pressure and critical voices call for greater accountability. Yet the very nature of global

excellence metrics makes rewarding excellence trickier in developing nations. How do we reward the careers of the best curious minds at lesser-known universities, who study issues of high regional importance that might have limited global consideration at that time?

With Africa steadily progressing in its science capacity, the anchoring of research excellence to such lists could render African-based academia a less attractive or less rewarding career choice, ultimately impeding the overall investment of science in the region.

I ask that Thomson Reuters and others take a more sophisticated approach that uses spatially explicit modelling (also known as individual- or agent-based modelling). Such models include extra dimensions — age and geography, for example — to highlight the best local outputs and people, and to account for their realized or *in situ* research impact.

Impact data need to be partitioned and not pooled globally. This would stop important scientists from developing regions falling through the cracks. In community-ecology analyses, we often standardize species-composition data to reduce the weight of the dominating species. This is important for observing the influence that the less-abundant species have on community patterns across the landscape. Impact assessment should do this too.

Locally sensitive lists would paint a less dismal picture of African research. This is more work for metrics institutions, but I feel that it is not unreasonable. (Some resources, such as the *Times Higher Education* World University Rankings, powered by Thomson Reuters, include structured sensitive metrics as regular features.)

I hope that a world that values the unravelling of a complex question, no matter how localized, has not disappeared. Is this not how the art of science originated, and what still draws curious scholars? I hope that there is still room to appreciate and respect local knowledge at a grand scale, on which global numbers do not exclusively dominate decision-making by funders, promotions by administrators, and the movement of young scientists. Otherwise it would be akin to judging happiness around the world on the basis of data on national gross domestic product.

If influential players in global science continue to propagate such crude centralized models, there would seem to be little hope for the development of a more equally distributed global science system, which in turn would drive a more equal and sustainable future. Such a system can be achieved only if we find ways to solve the problems in the regions we live in — independent of scale. ■

Casparus J. Crous is a researcher at the University of Pretoria in South Africa.
e-mail: casper.crous@fabi.up.ac.za

IMPACT DATA
NEED TO BE
PARTITIONED
TO STOP SCIENTISTS
FROM DEVELOPING
REGIONS FALLING
THROUGH
THE CRACKS.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/vcdpad

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

BIOCHEMISTRY

Marine enzyme can multi-task

An enzyme discovered in a marine bacterium catalyses complex reactions in a way that can be mimicked using simpler molecules, say biochemists at the Scripps Institution of Oceanography at the University of California, San Diego.

Bradley Moore and his team examined how a species of *Streptomyces* creates molecules known as merochlorins, which are promising antibiotics. They found that the bacterium uses a vanadium-dependent enzyme that first adds chlorine atoms to specific sites on a simple precursor molecule, then causes the precursor to wrap up — or cyclize — into the final merochlorin structure.

The researchers imitated the enzyme's unusual activity using a set of catalysts and small molecules to make previously overlooked products related to the initial merochlorins, in only five reaction steps.

Angew. Chem. Int. Ed. <http://doi.org/f2tm9p> (2014)

NEUROSCIENCE

Eyes offer window on dementia

The loss of cells in the retina might predict a common type of dementia years before behavioural changes set in.

A team led by Li Gan and Ari Green at the University of California, San Francisco, imaged the retinas of 12 people with mutations that cause frontotemporal dementia, a common cause of dementia in people under 60 years old. Even though most of the subjects had yet to develop symptoms, their retinas tended to be thinner than those of 24 people without the mutations. Mice lacking



GEOLOGY

Mobile rocks explained

Rocks that mysteriously slide across a dry lake bed are thought to be nudged along by large ice sheets. But previously, no one could fully explain how the rocks — some weighing more than 300 kilograms — scoot across Racetrack Playa in California.

Now a team led by Richard Norris at the Scripps Institution of Oceanography in La Jolla, California, has caught the rocks in action after tagging them with Global Positioning Satellite markers. On 20 December 2013, pools of ice in the lake started cracking in the morning sun. A gentle breeze then bumped 'windowpane' ice sheets against the rocks, moving them at a rate of 2–5 metres per minute. When the ice melted, more than 60 rocks had budged, leaving freshly formed trails (pictured) behind them. By the end of the winter, the farthest-moving rock had travelled 224 metres.

PLoS ONE 9, e105948 (2014)

the same gene also lost retinal neurons as they aged.

Retinal deterioration could be one of the earliest observable signs of frontotemporal dementia, the team says, and a useful biomarker in trials of drugs that slow the disease. *J. Exp. Med.* <http://doi.org/vdp> (2014)

GENOMICS

Ebola genomes decoded

Researchers from four countries have used deep sequencing to analyse the full suite of Ebola virus strains present in 78 people diagnosed

in Sierra Leone in the ongoing West African epidemic. The outbreak began in February 2014 in Guinea, and health officials had recorded 1,552 deaths as of 28 August, although this is thought to be well below the actual figure.

The study authors report that the current outbreak is caused by a virus variant that separated about ten years ago from viruses responsible for past Ebola outbreaks. This variant has accumulated 341 mutations since then, and 55 more as it spread among the people sampled in this study. In contrast to some past Ebola outbreaks, in which humans repeatedly acquired the virus from animal reservoirs, human-to-human transmission is driving the current outbreak. The researchers hope that the work will inform the design of Ebola diagnostics, vaccines and treatments.

Science <http://doi.org/vfk> (2014)

ANTHROPOLOGY

Sahara stopped human mixing

North Africa was a patchwork of human populations around the time our species trekked out of the continent.

Previous studies have proposed that the Sahara Desert was transformed into lush grassland during a wet spell between around 130,000 years and 75,000 years ago, offering few barriers to *Homo sapiens* expanding its range. However, researchers led by Eleanor Scerri at the University of Bordeaux, France, modelled the climate of northern Africa during this period and found that the Sahara, although much greener than today, still contained tracts of uninhabitable desert.

Comparisons of artefacts

MIKE HARTMANN

from 17 archaeological sites revealed geographic variability in the types of stone tool and techniques used to make them. The most similar tool assemblages tended to be those found closest to one another or those linked by green passageways.

Human populations in North Africa, including those that eventually colonized Europe and Asia, lived in semi-isolated groups that rarely mixed, the authors say.

Quat. Sci. Rev. 101, 207–216 (2014)

MATERIALS SCIENCE

Flexible solar cells work both ways

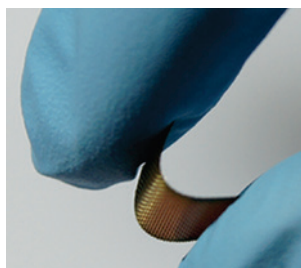
Material scientists have created a wearable fabric that can gather solar energy from either side.

Huisheng Peng at Fudan University in Shanghai and his colleagues made the material (pictured) by sandwiching

a textile, woven from metal fibres coated with a photoactive polymer, between two transparent and conductive carbon-nanotube sheets. The sandwich design means that the polymer solar cells can convert light to electricity regardless of whether they are illuminated from the top or bottom. This could make the fabric simpler to integrate into devices.

The cells convert just 1% of sunlight into electricity. But with efficiency improvements, the fabric could be used to power portable electronics, the authors say.

Angew. Chem. Int. Ed. <http://doi.org/f2tqbp> (2014)



SOCIAL SELECTION

Popular articles on social media

Fresh arguments over old data

Every once in a while, an academic argument turns into a social-media spectacle. A collision over a seemingly innocuous subject — genes and the pursuit of happiness — has attracted a large number of onlookers.

In 2013, a paper concluded that a person's approach to happiness can shape gene expression (B. L. Fredrickson *et al.* *Proc. Natl Acad. Sci. USA* 110, 13684–13689; 2013). The study and its first author, psychologist Barbara Fredrickson at the University of North Carolina in Chapel Hill, received much media attention at the time, but a new report by US and UK researchers claims that the findings were “artifacts of dubious analysis and erroneous methodology”. Stuart Ritchie, a human-intelligence researcher at the University of Edinburgh, UK, evidently enjoyed the show, as revealed by his tweet: “This demolition of a ‘genetics of wellbeing’ paper raised my wellbeing by a considerable amount.”

Proc. Natl Acad. Sci. USA <http://doi.org/vfp> (2014)



Based on data from altmetric.com. Altmetric is supported by Macmillan Science and Education, which owns Nature Publishing Group.

NATURE.COM
For more on popular papers:
go.nature.com/3ysfdp

SEVEN DAYS

The news in brief

POLICY

E-cigarette reports

Electronic cigarettes need more-stringent regulations, the World Health Organization (WHO) said on 26 August. The agency says that bans on indoor 'vaping' and on sales to minors should be put in place until there is more evidence of the devices' safety and the health risks are better understood. In a separate report, released on 25 August, the US Centers for Disease Control and Prevention found a threefold increase from 2011 to 2013 in the number of US school-age e-cigarette users who had not smoked previously. See page 24 for more.

Animal welfare

Labs certified by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International violate animal-research rules more often than non-accredited labs, according to a study by People for the Ethical Treatment of Animals (PETA). AAALAC International certification is the 'gold standard' for animal welfare; funding agencies including the US National Institutes of Health grant accredited labs privileges such as inspection waivers. PETA examined records for more than 800 US labs, and presented its findings on 25 August at the 9th World Congress on Alternatives and Animal Use in the Life Sciences.

Corals protected

On 27 August, the US National Oceanic and Atmospheric Administration listed 20 species of coral as threatened under the auspices of the Endangered Species Act. In November 2012, the agency had proposed listing

54 coral species as threatened and 12 as endangered. It had also suggested changing the status of the elkhorn (*Acropora palmata*) and staghorn (*Acropora cervicornis*) from threatened to endangered. The agency says that it narrowed the list owing to public comments as well as research findings on coral biology, habitat, distribution and abundance.

Reburied remains

A US federal court ruled on 27 August against scientists seeking to prevent the reburial of 9,000-year-old human remains. The bones were found on the campus of the University of California, San Diego, in 1976. In 2012, the university acquiesced to a request from the Kumeyaay Native American tribes to return the bones. But a group

of palaeoanthropologists sued the university in an attempt to block the return, claiming that the Kumeyaay were not culturally affiliated with the remains. The Ninth Circuit appeals court in San Francisco rejected the scientists' claim.

Data sharing

All genomics data generated through large-scale projects funded by the US National Institutes of Health (NIH) must be shared online, the agency said on 27 August. The NIH previously required such sharing only for genome-wide association studies. Researchers must also ask study participants for consent to share their data for use in future research — a policy that also applies to cell lines and clinical specimens such as tissue samples. See go.nature.com/hi4xsx for more.

Institute reform

The Japanese institute at the centre of a misconduct scandal has announced that it will be renamed and that the number of its laboratories will be cut. Although world renowned, the RIKEN Center for Developmental Biology in Kobe has been under fire after biochemist Haruko Obokata was found guilty of scientific misconduct. Its director Masatoshi Takeichi will also step down. See go.nature.com/fvyj36 for more.

EVENTS

Satellites stuck

A pair of satellites for Europe's Galileo global navigation system, which suffered a botched launch on 22 August, do not have enough fuel to reach their intended orbits, says the European Space



EGGERT JOHANNESON/AP

Seismic unrest at Bárðarbunga

Iceland's Bárðarbunga volcano erupted on 29 August, spurting lava out of a 1.5-kilometre-long fissure in a barren lava field (pictured). The surface eruption continued intermittently over the following days, and is the latest manifestation of a massive underground movement of magma,

which has created a buried wall of fresh rock some 45 kilometres long. As of 1 September, no ash had been emitted. But halfway around the world, the Rabaul volcano in Papua New Guinea spewed ash on 29 August, leading to the re-routing of several flights in the area.

Agency. The satellites, which are stuck in an elliptical orbit at a lower altitude than planned, might still be useful for some navigation or test purposes, the agency says. A report into the faulty launch is expected on 8 September.

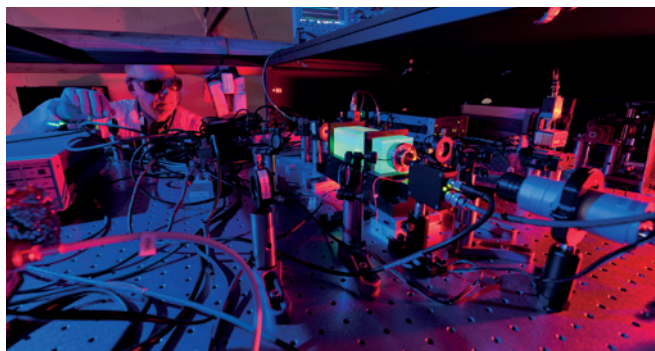
RESEARCH

Comet landing sites

The European Space Agency announced a list of candidate landing sites for its robotic comet probe, Philae, on 25 August. In the first ever attempt to land on a comet, the agency's Rosetta spacecraft will dispatch Philae to the surface of 67P/Churyumov-Gerasimenko in November. Studies have identified five potential sites on the rubber-duck-shaped comet — three on its 'head' and two on its 'body' — on the basis of landing constraints and scientific priorities. Sites will be assessed and ranked by 14 September.

Hologram test

An experiment designed to test whether the Universe is a two-dimensional hologram has begun collecting data, the Fermi National Accelerator Laboratory near Batavia, Illinois, announced on 26 August. The project scientists hypothesize that what we perceive to be a third dimension could actually be



information encoded in a two-dimensional reality. In principle, such a phenomenon would have a measurable signature, which the lab's Holographic Interferometer (pictured), or Holometer, will attempt to detect using twin lasers.

Mission reprieve

Two NASA missions considered on the verge of closure — the Mars Opportunity Rover and the Lunar Reconnaissance Orbiter — will continue for another two years, scientists close to the projects have confirmed. The news broke ahead of a long-awaited review to set priorities on NASA's planetary missions in the event that tight budgets force the agency to switch off operating spacecraft. The review is expected to be released on 3 September. See go.nature.com/1fi8m5 for more.

Ebola vaccine trials

The US National Institute of Allergy and Infectious Diseases announced on 28 August a series of phase I trials of Ebola vaccines in healthy volunteers. A trial of a vaccine developed with pharmaceutical giant GlaxoSmithKline will begin next week; another, developed by the Public Health Agency of Canada and licensed to the company NewLink Genetics in Ames, Iowa, is expected to undergo safety studies in the autumn. The trials will assess immune responses and vaccine safety. See page 13 for more.

BUSINESS

Space-junk system

US defence company Lockheed Martin in Bethesda, Maryland, has signed a deal with Electro Optic Systems in Canberra to build a space-debris tracking station in

Western Australia. The station will use advanced laser and optical technology to detect, follow and identify space junk that can damage satellites. It will complement radar-based systems such as the US Air Force's Space Fence, to be built by Lockheed Martin, which will track about 200,000 objects. There are an estimated 500,000 pieces of debris larger than a marble spinning around Earth at about 28,000 kilometres per hour.

Stem-cell trial

A landmark stem-cell trial that was abruptly stopped in 2011 has been given a green light to restart by the US Food and Drug Administration. Asterias Biotherapeutics of Menlo Park, California, announced on 27 August that in early 2015 it will start a safety and efficacy trial of a treatment involving nerve-cell precursors derived from human embryonic stem cells. Thirteen patients with spinal-cord damage in the neck will receive the therapy within a month of their injuries (see *Nature* 510, 18; 2014). Geron of Menlo Park, California, which originally started the trial, had stopped the study to focus on cancer treatments.

Investment ban

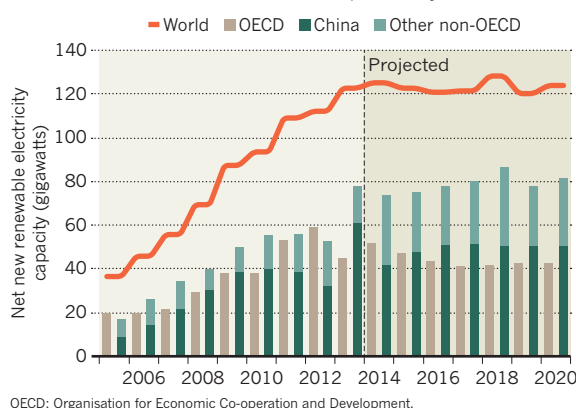
The University of Sydney in Australia has ceased to invest in coal mining, making it the first such institution in the country to do so, according to media reports. On 25 August the university announced that it will no longer invest in Whitehaven Coal, a Sydney-based company that is developing Australia's largest open coal mine. The following day, the institution said that its divestment extends to the whole coal and consumable fuels sector. The move follows a campaign from the environmental group Greenpeace, calling on the university to divest from Whitehaven Coal.

TREND WATCH

Renewable-power capacity grew faster than ever before in 2013 — but growth should level off by 2020, the Paris-based International Energy Agency says in this year's *Renewable Energy: Medium-Term Market Report*. The report, released on 28 August, says that policy and market risks are making it difficult to forecast deployment, including doubts about how to integrate renewable power into the electricity grid; and uncertainty over whether subsidies and tax incentives will be maintained.

FLAT PREDICTION FOR RENEWABLES

The International Energy Agency expects most of the growth in capacity to come from countries outside the OECD, particularly from China.



NATURE.COM

For daily news updates see:
www.nature.com/news

NEWS IN FOCUS

OCEANOGRAPHY Researchers focus on prospective El Niño in bid to bolster climate models **p.15**

NATURAL HAZARDS Swiss authorities brace for Alpine lake tsunami **p.16**

SEISMOLOGY California revisits hazard-warning system in wake of quake **p.18**

CAREERS Why star students choose to abandon the lab **p.20**



The Ebola outbreak, which has killed more than 1,500 people, has devastated the impoverished countries of West Africa.

INFECTIOUS DISEASE

Ebola drug trials set to begin amid crisis

Testing drugs in the middle of deadly disease outbreak is challenging but can be done.

BY DECLAN BUTLER

William Pooley, a nurse from the United Kingdom who caught Ebola while treating patients in Sierra Leone, is the latest person in the current outbreak to receive the experimental drug ZMapp. Yet there is still no evidence that the drug, or any of the other treatments under development for Ebola, are safe or effective in people. To change this state of affairs, the World Health Organization (WHO) is this week con-

vening a major meeting to prioritize the most promising experimental Ebola drugs and vaccines, and to discuss how clinical studies can best be carried out in the chaotic environment of a deadly disease epidemic.

Although the drug's manufacturer Mapp Biopharmaceutical of San Diego, California, has given out doses of ZMapp to seven people so far, this was authorized on a case-by-case, 'compassionate use' basis. No drugs or vaccines for Ebola have been approved by regulators.

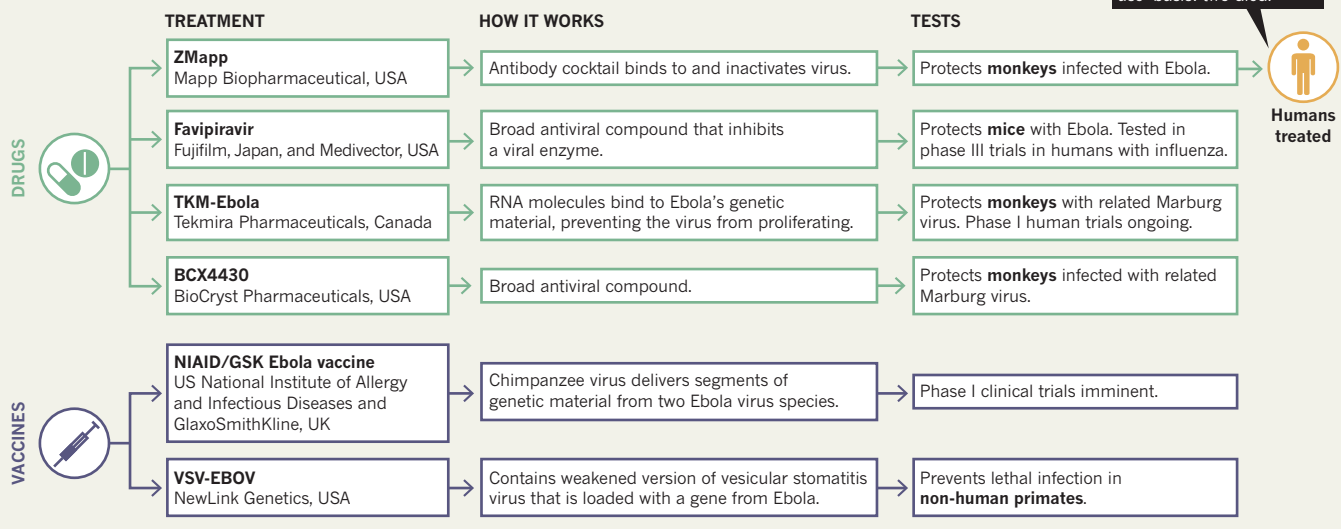
The first phase of clinical trials, to test for

a product's safety, is usually carried out in healthy volunteers in facilities with sophisticated clinical-trials infrastructure. But an unusual combination of factors — the difficulty of implementing public-health measures to control the disease's spread in the affected countries, the huge social and economic disruption that it is causing and the fact that the current outbreak kills about 53% of the people it infects — makes this crisis an exception.

Last month, an expert panel convened by the WHO concluded unanimously ►

DRUGS AND VACCINES TO WATCH

This week the World Health Organization is convening a meeting to discuss Ebola drugs and vaccines that might be prioritized for trials during the current outbreak. Here are some that delegates will consider.



► that it is ethical to use unapproved drugs and vaccines in this outbreak, provided that every effort is made to gather the best scientific data on their safety and efficacy. And on 4–5 September, 150 Ebola scientists, industry executives, clinical-trials experts, ethicists and regulatory officials are due to meet at the WHO's headquarters in Geneva, Switzerland, to identify and prioritize the most promising products for use in clinical trials (see 'Drugs and vaccines to watch').

Although the slow regulatory route for testing drugs and vaccines is ill-adapted to the rapid response needed during an outbreak, it is an obstacle that must be overcome, says Jeremy Farrar, a tropical-disease specialist and head of the London-based biomedical-research charity the Wellcome Trust. "We don't have the luxury of being able to test these in humans outside of an epidemic setting."

Large and complex trials, in which treatments are allocated at random to a homogeneous group of people, would be impractical in the field, say infectious-disease researchers: the countries affected by the Ebola outbreak have fragile health-care systems that have been overwhelmed by the outbreak and a health-care workforce ravaged by the virus.

The next steps are therefore likely to be the introduction of small-scale trials for both safety and efficacy, in which people with Ebola receive the drugs in an organized manner, and the outcomes are systematically recorded, says Daniel Bausch at Tulane University in New Orleans, Louisiana. Ebola's high mortality rate also means that if a drug works, this should

be easy to detect even in trials involving just a few dozen patients.

At the same time, doctors and clinicians must better structure the data

obtained from cases of compassionate use to increase their scientific value, says Bausch, who is one of the world's leading experts in fighting Ebola outbreaks and has been caring for patients in Sierra Leone.

Pooling results from different small trials and from compassionate use — perhaps through a central coordinating body such as the WHO — will be key to extracting the most useful data, says Farrar. That will require the various research organizations, charities and drug companies providing the medications to agree to standardize how they collect their data, for example monitoring a patient's viral load and recording how long patients had been infected when they first received the drug. "Let's say five different interventions get tried in the next three months," says Farrar. "If each of those gathers completely different, non-comparable information, we are likely to end up with no conclusive information."

A coordinated approach could yield data that, although not as high quality as those obtained from gold-standard randomized clinical trials, would be good enough to enable decisions to be made on potential treatments, says Farrar. Sharing results could also allow an iterative process in which ineffective products are dropped, and effective ones scaled up.

But implementing even slimmed-down trials will be possible only at the best treatment centres, because the logistics elsewhere make any clinical work almost impossible, points out Bausch. "In Liberia right now, it's unrealistic to think there is any type of infrastructure; they

are struggling to have enough beds" for Ebola patients, he says. In an outbreak situation, even simple matters such as recording data for a trial become difficult, he adds. "Filling out a form on 300 patients when you have just two nurses and a doctor is just not realistic."

The WHO meeting will also discuss vaccines. Because these are designed for people not infected with the virus, the ethical equation for testing vaccines is different.

On 28 August, Anthony Fauci, director of the US National Institute of Allergy and Infectious Diseases, announced that the agency is working with vaccine makers and an international research consortium, including the Wellcome Trust, to conduct phase I safety trials of two candidate vaccines. The trials will be carried out in healthy volunteers in the United States, the United Kingdom, Gambia and Mali, all of which have good clinical-trials infrastructure, before being tested or introduced into outbreak areas.

Getting any products out in time to be of significant help in controlling the outbreak will be challenging, says Bausch. Still, the investment could help to control future Ebola outbreaks.

Robert Fowler, a critical-care specialist at the University of Toronto, Canada, who has been working for the WHO in Guinea and Sierra Leone, says that although he is supportive of the testing of new drugs and vaccines, improvements in basic treatment should not be overlooked. Fowler is preparing to publish results showing that providing intravenous rehydration to restore electrolyte balance can greatly reduce mortality rates. Such supportive care is available at only a few treatment centres in the region, but could help to reduce mortality right now, he says. ■

Additional reporting by Nicole Skinner.

➔ **NATURE.COM**
For more on the Ebola outbreak in West Africa see:
go.nature.com/m55ual



The strong El Niño weather event of 1997–98 caused flooding and massive damages in California.

OCEANOGRAPHY

Stalled El Niño poised to resurge

Studies of brewing weather event test understanding of past and future climate.

BY MARK ZASTROW

When Julia Baum tried to fly to the remote Pacific atoll of Kiritimati with a chest freezer in her luggage last month, her airline objected. Baum needed it to store samples of the algae that underpin the atoll's fragile coral ecosystem. "It was a lot of smiling and begging," she recalls. "It's for science! It's for coral reefs! Help us!" Airline officials let it go.

This year Baum, a marine biologist at the University of Victoria in Canada, has an added incentive to head into the field: the weather pattern known as El Niño is brewing, and she wants to see what the warming it brings to the tropical Pacific Ocean will do to the algae that live in the reefs of Kiritimati, part of the island nation of Kiribati.

Hers will be one of many teams watching this potential El Niño. A growing array of satellites, moored and free-floating buoys and autonomous underwater vehicles is tracking changes in ocean temperatures and atmospheric conditions. Those data are flowing into

climate models that try to predict the strength and timing of this year's El Niño event — the erratic behaviour of which has already tested forecasters. If researchers could learn to predict such events with precision, it would significantly improve their understanding of the past and future climate.

The first signs of an El Niño came in January, when the east-to-west trade winds that blow across the tropics suddenly weakened, and a burst of winds from the west triggered a slow surge of warm water into the eastern equatorial Pacific. In 1997, similar conditions helped to set in motion one of the strongest El Niño events on record, which caused extreme rainfall along the western coasts of North and South America and drought in Australia and southeast Asia, resulting in thousands of deaths and tens of billions of dollars' worth of damage.

Sustaining an El Niño requires the ocean and the atmosphere to work together. Normally, the warming in the eastern Pacific that presages such an event strengthens wind patterns that push even more warm water eastward. But this year, the atmosphere did not play its part.

As a result, the ocean cooled in May, June and July, and the El Niño stalled.

Anthony Barnston, a seasonal climate forecaster at Columbia University in New York City, says the latest data show that winds from the west are again forming — a second chance for a full-blown El Niño. His team estimates that there is a 75% chance that a weak to moderate event will form by the end of this year, just a bit later than researchers had thought. The predictions are in line with those from the US National Oceanic and Atmospheric Administration (NOAA), which puts the chance of a weak to moderate El Niño at 65%.

Whatever transpires, the next few months will be an important test for the latest generation of seasonal forecast models. Scientists struggled to interpret the wildly divergent predictions provided by models before the El Niño of 1997–98, but this time there is cause for optimism. When the warming of the eastern Pacific started to slow this summer, the latest models — which run at higher resolution and on more powerful computers than their predecessors — predicted that El Niño would stall in the summer and then resurge. "If we don't see an El Niño this year, then we have a big black mark on the model performance," says Barnston.

Other scientists are trying to learn what they can about El Niño's links to future climate change.

Climate models disagree on whether global warming will alter El Niño's strength or frequency. But even if the general pattern of El Niño events does not change, the warming produced by an individual El Niño in an overall warmer climate could generate more extreme weather, says Wenju Cai, a climate scientist at the Commonwealth Scientific and Industrial Research Organisation's Marine and Atmospheric Research centre in Aspendale, Australia. "The world is not waking up to that," says Cai, whose modelling work suggests that by the end of this century, El Niño events could become twice as common as they were for most of the twentieth century (W. Cai *et al.* *Nature Clim. Change* 4, 111–116; 2014).

Back in the tropical Pacific, Baum's colleagues hope to gain a fuller understanding of El Niño's history by analysing the oxygen content, salinity and temperature of water samples from Kiritimati and the Palmyra atoll, a US territory roughly 680 kilometres away. Researchers will use the data to help to calibrate records of past climate that are preserved in fossilized coral. Shifting ratios of oxygen isotopes captured in coral layers can reveal changes in ocean temperature, providing a record of El Niño events going back thousands of years.

One thing is certain, says Michael McPhaden, an oceanographer with NOAA in Seattle, Washington: however this year's El Niño watch plays out, it will influence research for several years. "Why has nature surprised us in this startling way?" he asks. "Figuring it out is going to be really important." ■



ALAMY/PRISMA BILDAGENTUR

The area surrounding Lake Lucerne in Switzerland experiences a magnitude-6 earthquake about once every 1,000 years.

NATURAL HAZARDS

Switzerland braces for Alpine lake tsunami

Authorities in Swiss canton blaze a trail by factoring the risk into hazard planning.

BY LAURA SPINNEY

The land of chocolate and clocks could soon be known for something quite different: tsunamis. Authorities in Nidwalden, a canton in landlocked Switzerland, are factoring the risk of a tsunami in Lake Lucerne into their hazard plans. It is the first official acknowledgement of such a threat in Europe's Alpine region — and comes in step with findings that the risk of tsunamis in the area, which is home to around 13 million people, is much higher than previously thought.

Most tsunamis occur in the ocean but they can also occur in enclosed bodies of water, when underwater sediments shift as a result of an earthquake, falling rocks or underlying instability. The hazard that such events pose is outsized. “The same source placed inside a lake can have a bigger impact than along the coast of an open ocean,” says Hermann Fritz, who studies tsunamis at the Georgia Institute of Technology in Atlanta. A tsunami at Lake Yanawayin in Peru in 1971 is thought to have killed 400–600 people.

Although not as seismically active as Peru or Japan, the Alps do experience earthquakes from time to time: one as strong as magnitude 6 occurs roughly every millennium around Lake Lucerne, for instance.

The issue of tsunamis in Alpine lakes grabbed the spotlight two years ago when limnogeologist Katrina Kremer, then at the University of Geneva, Switzerland, and her colleagues reported evidence for a major tsunami in Lake Geneva in AD 563 that had wiped out communities living on its shores¹.

Kremer, who has since moved to the Swiss Federal Institute of Technology (ETH) in Zurich, has upped the count to five probable tsunamis in Lake Geneva over the past 4,000 years, including another devastating one during the Bronze Age around 3,400 years ago². Her team relied on cores taken by drilling 30 metres into the lakebed. The cores represent 4,000 years' worth of sediment deposition, and show telltale structures of lakebed displacement that could have caused a tsunami. Kremer presented the work on 18 August at a meeting of the International Association of Sedimentologists in Geneva.

The AD 563 event, the largest of the five, occurred when part of a mountain fell onto an unstable underwater delta of the Rhône river, which flows into Lake Geneva (see ‘Origin of a wave’). The falling rock forced the collapse of

sections of the delta, creating a wave that was 8 metres high when it hit and wiped out the old city of Geneva at the other end of the lake.

The Bronze Age tsunami was 6 metres high, and may provide another example of tsunami-related devastation in the Lake Geneva region, because it fits with a possible abandonment by lake-shore dwellers. “We think this could explain an occupation gap in early Bronze Age sites on the northern shore of the lake,” says Kremer. Albert Hafner, an archaeologist at the University of Bern, finds the hypothesis convincing: “They left their sites for many reasons, but tsunamis may have been one of them.”

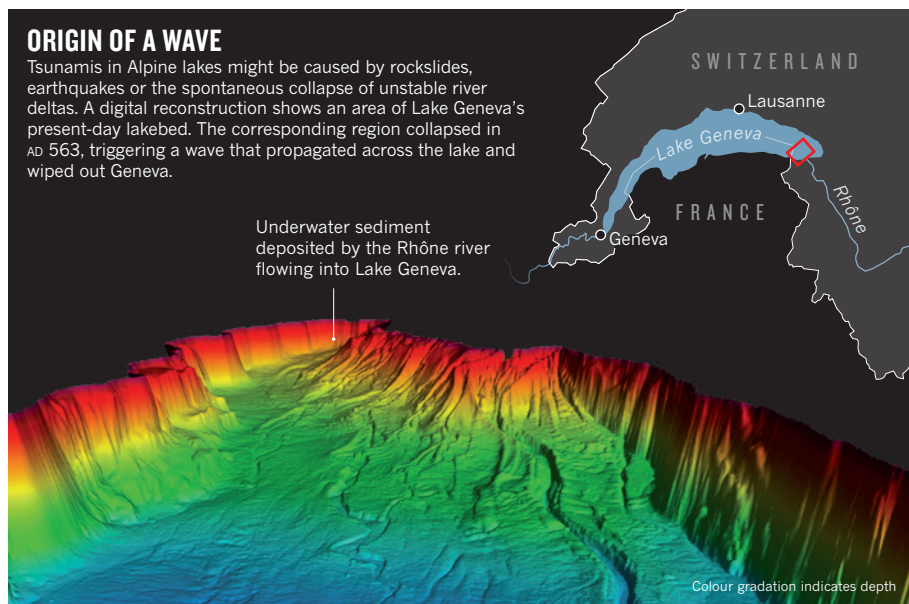
The episodes add to growing evidence for historical tsunamis in other Alpine lakes, including two events³ in Lake Como in northern Italy, in the sixth and twelfth centuries AD, and one in Lake Bourget in the French Alps in 1822. A witness of the 1822 tsunami described the water as “boiling like soup”, an effect caused by the release of methane when an earthquake stirs up organic debris in the lake.

In the past decade, a history of tsunamis has also emerged for Lake Lucerne. Research by Flavio Anselmetti, a geologist at the University of Bern, and his colleagues reveals⁴ that the lake experienced two tsunamis in the seventeenth century. The first, in 1601, reached 5 metres in

“Taking such risks into account is quite complicated.”

ORIGIN OF A WAVE

Tsunamis in Alpine lakes might be caused by rockslides, earthquakes or the spontaneous collapse of unstable river deltas. A digital reconstruction shows an area of Lake Geneva's present-day lakebed. The corresponding region collapsed in AD 563, triggering a wave that propagated across the lake and wiped out Geneva.



height. The amount of sediment on some of the lake's underwater slopes has since increased, and a strong earthquake could easily dislodge it, triggering a tsunami, says Anselmetti.

Nidwalden, which borders the lake, has now commissioned Anselmetti's team to assess the likelihood of such an event for different scenarios, including earthquakes of various strengths or a rockslide — as well as the probable effect on the surrounding land. Aided by geological data and computer simulations, they plan to create maps of the inundation in each scenario. The authorities will use the charts to draw up evacuation plans and to guide future construction.

Although the risk of a tsunami in Lake Lucerne seems remote, the possibility is real enough that insurers might want to know about it, says Anselmetti. Underwriters typically consider events that might occur within the next 500 years worth insuring against, and it has now been more than 400 years since the most recent magnitude-6 earthquake at Lake Lucerne.

Nidwalden is, for now, blazing a lonely trail with such plans. Conditions differ across regions, provoking varying reactions to the threat of Alpine tsunamis. With many more people living on its shores and different physical and geographical factors at play, a tsunami in Lake Geneva could have more devastating effects than one in Lake Lucerne, but is less

likely to occur in the next 500 years. "Taking such risks into account is quite complicated," says Jacques Martelain, geologist for the canton of Geneva. "What should we do, knock the city down and rebuild it elsewhere? We could take measures to protect the built-up lake shore, but that seems disproportionate."

Other regions may yet follow in Nidwalden's footsteps. Kremer's latest work shows that smaller, but still damaging, tsunamis occur more frequently than bigger ones. When it comes to crucial infrastructure, insurers deem periods longer than 500 years worth planning for: an event that poses a threat to nuclear-power plants, for instance, need recur only every 10,000 years to be considered.

Lives can be saved if people are trained to respond to a tsunami, but until earthquake prediction improves, planning efforts will not be able to mitigate damage completely. "The population," says Fritz, "must learn to live with the hazard to some extent." ■

1. Kremer, K., Simpson, G. & Girardclos, S. *Nature Geosci.* **5**, 756–757 (2012).
2. Kremer, K. et al. *Earth Planet. Sci. Lett.* **385**, 28–39 (2014).
3. Fanetti, D., Anselmetti, F. S., Chapron, E., Sturm, M. & Vezzoli, L. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **259**, 323–340 (2008).
4. Schnellmann, M., Anselmetti, F. S., Giardini, D. & McKenzie, J. A. *Eclogae Geol. Helv.* **99**, 409–428 (2006).



Q&A



Ebola doctor on patients he cured in the United States
go.nature.com/ocqdbu

MORE NEWS

- Publication bias afflicts social sciences go.nature.com/vlrfsc
- Trans-Atlantic row over stellar distances is reignited go.nature.com/k4mprd
- Portable NMR test for malaria shows diagnostic promise go.nature.com/jh9xkl

SEISMOLOGY

California quake puts warning system in the spotlight

Scientists, politicians and business leaders will discuss practicalities of a regional network.

BY ALEXANDRA WITZE

The Bay Area Rapid Transit (BART) trains of San Francisco were not running at 3.20 a.m. on 24 August when a magnitude-6 earthquake hit the region. But the BART computers were working — and they knew what was about to happen 10 seconds before the ground started shaking, thanks to an alert from California's prototype warning system. Had the quake struck in the middle of the day, the computers would have instructed the trains to slow to a stop — potentially avoiding a derailment and saving the lives of passengers.

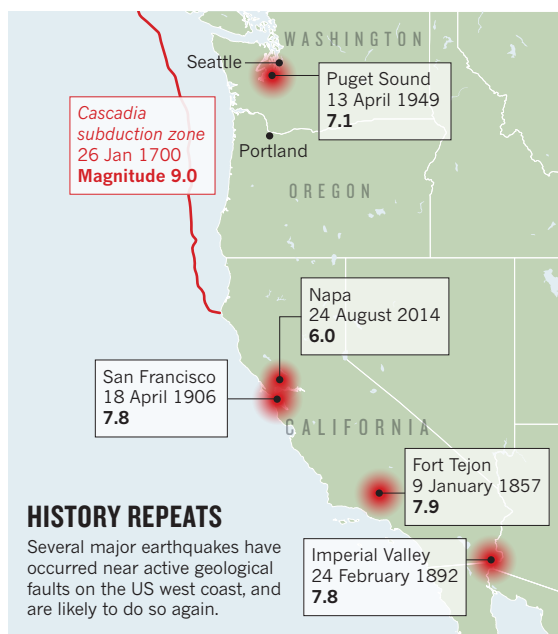
With such a scenario in mind, researchers are pushing hard to wire a much broader swathe of the US west coast, from California to Washington, with a fully fledged earthquake warning system. The US\$120-million project has not yet secured funding, but the latest quake, centred near Napa, California, which caused several hundred million dollars' worth of damage, could tip political fortunes in its favour. Seismologists and emergency-services managers will discuss the practicalities of implementing such systems at a conference starting on 3 September at the University of California, Berkeley.

"I'm hoping the Napa quake has created a sense of urgency," says Alex Padilla, a California state senator who represents a district near Los Angeles. Padilla introduced a bill that Governor Jerry Brown signed into law last September, which gives the state until 2016 to set up an early-warning system before the legislation expires.

Yet the law does not specify where the \$80 million for California's part of the system would come from. That puts US seismologists at a crucial turning point: if they are unable to push through an early-warning system in time, they might struggle to get one at all.

Several countries, including Mexico and Japan, have extensive warning systems in place to protect their key infrastructures. And in Istanbul, seismic sensors allow one of Turkey's main natural-gas providers to shut down its pipelines if the ground is about to start shaking. Similar warnings flow to nuclear power plants in Romania.

California has the backbone of a warning



system — ShakeAlert, a project run by a consortium of universities and the US Geological Survey (USGS). The system uses real-time information from seismic waves arriving at California's sensor network to send warnings that secondary, more damaging waves are on their way. The alerts go to scientists as well as to about 150 organizations, emergency managers and other testers.

And the system works. During the recent Napa quake (see 'History repeats'), ShakeAlert provided about 10 seconds of warning to San Francisco, which is about 50 kilometres from the epicentre. "If this isn't a sufficient demonstration of the system to fund full-blown implementation, how big does a quake have to be?" asks Richard Allen, a seismologist at the University of California, Berkeley.

ShakeAlert was developed with \$6.5 million from the philanthropic Gordon and Betty Moore Foundation in Palo Alto, California, but that money will run out next year. So far, no one has stepped up with the funds to turn it into a mature operational system. In an ideal world, the USGS would have set up a federally run warning system years ago, says Thomas Heaton, a seismologist at the California Institute of Technology in Pasadena.

But USGS budgets have been shrinking for years, and states will probably end up footing

at least part of the bill. To Heaton's mind, it is not an ideal solution because states do not necessarily have the resources or expertise: having a state in charge is "like Florida running a hurricane warning system", he says.

Padilla and other politicians have been trying to secure the money through different infrastructure projects to upgrade California's current seismic network and add hundreds of new stations. An earthquake early-warning system will never be perfect: it will always have a blind zone, an area too close to the quake to receive a warning. But that zone can be shrunk by adding stations and improving the network, Allen says.

For example, the decade-old sensor at the seismic station closest to the Napa quake's epicentre took 2.8 seconds to register the shaking, process it and relay it to the network that monitors Californian earthquakes. Had the sensor been newer, it could have done all that in just 0.3 seconds — fast enough to alert the town of Napa itself that a quake was imminent, allowing residents to dive for cover, Allen says.

To be most effective, the warning network would also need to link to infrastructure, as it does to the BART computers. "We're really talking about a modern command-and-control system," Heaton says. "To turn our seismic networks into something that's automatically controlling trains and elevators and the like, we need a team of software designers working for several years," he explains.

Such connectivity becomes even more important were California to integrate its system with seismic networks farther north in Oregon and Washington. There, the offshore Cascadia fault zone has potential to let loose a quake as big as magnitude 9. The fault is relatively far from cities, so an early-warning system might provide minutes of warning to Portland, Seattle or Vancouver.

Expanding into Oregon and Washington would require \$40 million over five years. Seismologists in the region aim to trial alerts using testers — much as ShakeAlert is doing — by next spring, says John Vidale, a seismologist at the University of Washington in Seattle.

"We don't want to wait until we have a major damaging earthquake with hundreds of fatalities," says Allen. "We want to build it now." ■

SOURCE: USGS



THE ONES WHO GOT AWAY

Sometimes, the brightest stars in science decide to leave.
Nature finds out where they go.

BY EWEN CALLAWAY

When Soroosh Shambayati left his organic-chemistry lab, he didn't leave chemical synthesis behind. As a chemist PhD turned investment banker, he started working in the derivatives market in the 1990s. The transactions involved arranging a complex series of trades in a precise order, and it reminded him of synthesizing an organic compound, reaction by reaction.

As a graduate student, Shambayati had excelled at synthesis, just as he did at

everything he turned his hand to. He was "other-worldly brilliant", says his former adviser Stuart Schreiber. He juggled three distinct projects during his PhD, one in organic synthesis, one in theoretical physical chemistry and a third in biochemistry and immunology. He was also calm, thoughtful and well read: his bookshelf spans science philosophy, evolutionary biology and physics. Schreiber, a biochemist at the Broad Institute in Cambridge, Massachusetts, knew that if

ILLUSTRATIONS BY SEÑOR SALME

Shambayati wanted to become an academic scientist, he was sure to succeed. “It was very clear to me that he was going to become a star,” he says. But Shambayati chose the financial world — and excelled there instead: he is now chief executive at Guggenheim Investment Advisors (Suisse) in Geneva, Switzerland, a firm that manages billions of dollars for wealthy families and foundations.

Shambayati is among the hundreds of thousands of scientists who train in academia but then leave to follow a different career. According to the latest survey of doctorate recipients conducted by the US National Science Foundation, nearly one-fifth of employed people with science and engineering PhDs were no longer working in science in 2010. This is partly due to a lack of room at the top. In the United States, the number of PhDs entering the workforce has skyrocketed but the number of stable academic jobs has not. In 1973, nearly 90% of US PhDs working in academia held full-time faculty positions, compared with about 75% in 2010.

A common perception is that the weaker science students are forced out of a competitive field, leaving the brightest stars to secure the desirable academic positions. But as Shambayati's story shows — and as most mentors know — this is not the full picture: sometimes the scientists who move on are the ones with the most promise. Their motivations are diverse: some want more money, or more time with family; others are lured by opportunities elsewhere. To get a better sense of why talented scientists are leaving academia and how their training influences their lives, *Nature* contacted group leaders recognized for mentoring and asked: “Who was the one who got away?”

FROM CHEMIST TO CAPITALIST

Shambayati was born in Iran, attended school in Sweden, and then won a scholarship to study chemistry and mathematics at a university in Los Angeles. As an undergraduate, he was drawn to science for its pursuit of objective truth and the opportunity for discovery. A PhD was the obvious next step, and he found a perfect fit in Schreiber's lab in the late 1980s. “When I met Stuart, it was almost a sense of meeting the da Vinci of what I had in my head as science,” Shambayati says. At the time, the lab was focused on synthetic organic chemistry, but Schreiber also took risks — such as branching into biology — which meant that Shambayati could pursue his broad, three-pronged PhD.

The realities of doing science, however, soon butted heads with Shambayati's idyllic view of it. He found that chemical synthesis was slow and full of setbacks — “a bit like banging your head against the wall for long periods of time,” he says — and he was put off by the political aspects of science, exemplified at the time by bickering over who discovered HIV. Still, he did not hesitate to apply for

“I DON'T THINK I'M THE MOST NATURAL BANKER OR FINANCIAL THINKER, BUT I KNOW THAT I HAVE A CERTAIN AMOUNT OF INTELLECTUAL ABILITY.”

Soroosh Shambayati



faculty jobs at several top universities while finishing his PhD, and he received more than one offer.

While in New York for a job interview at Columbia University, Shambayati met up with a friend in banking, who was shocked to learn how little an assistant professor earned. He encouraged Shambayati to move into the financial world instead. “I said ‘you're nuts,’” Shambayati remembers. “I know nothing about banks or banking. Why would anybody want to interview me?”

But a good salary was tempting to Shambayati, who felt a deep obligation to support his family; his parents had fled Iran after the 1979 revolution, leaving their house and savings behind. Shambayati set up an interview with his friend's bosses at Banker's Trust, which was later bought by Deutsche Bank. The investment bank was a leader in derivatives trading and was looking for quantitative, analytical thinkers such as Shambayati. He accepted a job earning many multiples of an academic salary, figuring that he could always go back to do a postdoc if things did not work out.

They did. Finance was an eye-opener for Shambayati, who worked on chaotic emerging markets, losing and making back tens of millions of dollars in a day. He found the trading floor, surrounded by his colleagues, not unlike his chemistry lab, but with “even less privacy,” he says. His career progressed quickly and he moved on to jobs at Goldman Sachs, Citigroup and then Lehman Brothers.

(He was there in 2008 when the firm abruptly went into bankruptcy, catalysing the global economic crisis — an experience he compares to “being in a plane crash”).

Schreiber never questioned Shambayati's decision to leave science. He works hard not to presume that his best students will follow in his footsteps, he says. And Shambayati credits some of his success to the influence of his former mentor's approach to science, based on calculated risks, hard work and creativity. “I don't think I'm the most natural banker or financial thinker, but I know that I have a certain amount of intellectual ability that I can utilize, if I work very hard,” he says. “That was very much informed by the way Stuart operates.”

FROM PHYSICS TO CYBERSPACE

Sometimes, the decision to leave science is partly push, partly pull. That was the case for Renata Sarno, who after eight years in theoretical physics ran up against a dearth of academic jobs in her native Italy. Then the World Wide Web happened. Sarno started an online business — one of Italy's first — that would eventually sell for hundreds of millions of euros.

“She was a very brilliant student,” remembers her supervisor Giorgio Parisi, a theoretical physicist at the University of Rome. She could comprehend problems posed in theoretical physics and then determine how a computer might tackle the challenge. Sarno came to Parisi's lab to finish her undergraduate thesis in mathematical physics in 1987, and continued to work with him through a PhD and postdoc position. She helped to build a supercomputer, one of the world's fastest at the time, then used it to model subatomic particles called fermions using lattice gauge theory, which divides



“I FOUND MYSELF WITHOUT A SALARY AND A BIG ROAD FULL OF POSSIBILITIES IN FRONT OF ME.”

Renata Sarno

continuous space-time into series of discrete points. She was inspired by Parisi's diverse interests, such as protein folding and neural networking, as well as his desire to tackle new computational problems in particle physics.

An academic life seemed ideal to Sarno, and Parisi thought that she would make an excellent group leader. But when Sarno's postdoc funding ran out in 1994, she was unable to find more money, and there were few job opportunities for those starting out in science. Women were, and are, exceedingly rare at the highest levels of physics in Italy, and she also felt that discrimination held her back.

But there were opportunities opening up elsewhere. A year earlier, the European particle-physics lab CERN had made the World Wide Web public, and Sarno saw a chance for people with a background in computer science. "I found myself without a salary and a big road full of possibilities in front of me," she says. "I decided to take this road."

With three colleagues and roughly €10,000 (US\$13,000), Sarno launched a series of websites, including a travel website called Venere that was one of the first to offer reservations for hotels and others services. The team created tools to allow people to discuss and rate hotels, a novelty at the time. Sarno says that her research on problem solving and computation was good preparation for running a dot-com business, but she was also inspired by Parisi's tutelage: "My choice to go to the Internet was, in some sense, a choice to do something that is big from scratch, in a direction where nothing exists." Parisi was sorry to see one of his star trainees leave the lab, but he is proud of her success. Sarno and her colleagues sold Venere to the online travel company Expedia for around €200 million in 2008.

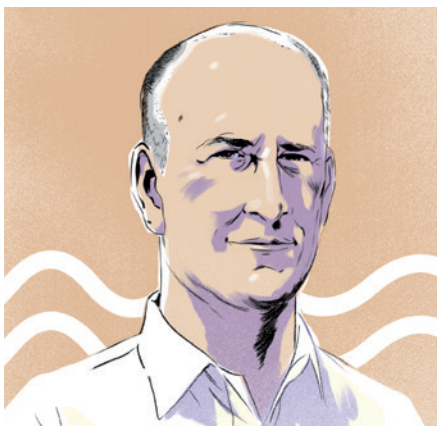
Sarno has not completely let go of research. After selling off Venere, she started a foundation to support research into blue-cone monochromacy, a rare genetic disease that runs in her family and which causes vision problems and colour blindness. She stays in touch with Parisi: when she wanted to know how different photoreceptors were arranged in the human retina, she sought his help in interpreting micrographs, and she is hopeful that delivering genes to the retina could help to cure the disease. Just like particle physics and the Internet, she sees an opportunity in gene therapy to solve problems in a new field.

FROM PHYSIOLOGY TO STAY-AT-HOME DAD

When Eric Pane started a PhD in physiology, he was already onto his second career — or was it his third? He had been teaching at a prestigious primary school in California, which was filled with the offspring of ambitious, demanding parents. Meanwhile, he was simultaneously working as a waiter, tutor and babysitter to make ends meet. The heavy workload was wearing him down, and he wanted a way out. "I didn't want to die in a

"I COULD HAVE GONE FOR IT, BUT OUR CHILDREN WOULD HAVE BEEN RAISED BY WOLVES AND GYPSIES AT THAT POINT."

Eric Pane



kindergarten classroom of a heart attack at age 70," he says. So Pane went back to university to study biology. He found that fish physiology, in particular, fascinated him for its efforts to understand how a complex living organism was put together. "I liked the 'take apart the toaster, take apart the fish' aspect," he says.

As Pane read the scientific literature, he repeatedly noticed the name of Chris Wood, a physiologist at McMaster University in Hamilton, Canada. Pane contacted him, and asked to join his lab. "He was sort of an unusual applicant," Wood recalls. Most of Wood's graduate students arrived just out of their undergraduate degrees, but Pane was already 31 when he joined the lab in 2000 and embarked on a PhD.

Pane quickly made an impression. He investigated the toxic effects of nickel on fish and became the lab's most productive member, publishing seven papers as part of his thesis, collaborating on another three and winning an industry award for his research. "He was one of those dream graduate students that everybody wishes they had. They come to you and generate their own ideas," says Wood. "I figured this guy was going to become a faculty member some day. He was the whole package."

With that faculty job in mind, Pane moved on to a postdoc studying ocean acidification at the Monterey Bay Aquarium Research Institute in Moss Landing, California. But when it came to the next career move, the rest of life had intervened. By then, Pane and his wife Michiko had had two boys, Michiko's career as a research-grant administrator was flourishing, and they had settled in the San Francisco Bay area. While Pane's colleagues were being interviewed for faculty positions all over the world, he was limited to local

universities so that the family could stay put. The time demands and comparatively low pay of an assistant professor also weighed heavily: Michiko worked 60-hour-plus weeks, and Pane wanted to be able to pick up his boys from school. "I could have gone for it," he says, of the faculty job, "but our children would have been raised by wolves and gypsies at that point." Pane now teaches two days a week at a community college — much of it done online. He spends the rest of his time as a stay-at-home parent.

Pane is still envious of the environment of productivity and achievement his former mentor encouraged, and he tries to build that culture in his own classes. He is proud to be described as "tough, but fair" on the website ratemyprofessor.com.

Wood was surprised that his former student took an alternative career route. "Initially I was disappointed. But it's up to everybody to make their own path in life," he says. Wood's reaction mirrors those of many group leaders who see their star students quit the lab: pleasure that he or she has found a life they are happy with, tinged with regret that science will not benefit from the student's talent. But most of the scientists *Nature* contacted while researching this story saw mentoring as more than simply grooming a new generation of professors, and they recognize that non-academic jobs are a good application of scientific training.

In a paper published in 2012, a team of science-policy researchers tried to work out why scientists exit academia (B. van Balen *et al. High. Educ. Policy* 25, 313–334; 2012). The team, led by Peter van den Besselaar at VU University Amsterdam, compared 21 pairs of Dutch researchers who were of comparable ages and working in similar fields. Both were deemed very talented in their early careers, yet one left academia and the other stayed. The team found few concrete differences between stayers and leavers to explain their divergent paths; both published similar numbers of papers that were cited just as highly, for example. But the stayers were more likely than the leavers to have had a stimulating mentor, support from a partner and good job opportunities. When it comes to securing an academic job, says van den Besselaar, "it may be a question of luck — are you there at the right moment and the right place?"

Nearly a decade out of Wood's lab, Pane still misses it — and he wonders every day if he made the right choice. "It's just a fork in the road. You've got to go one way or the other, and you're always going to regret what the other one looked like." But on this particular day, a Friday in early July, there is no time for regret because he has got plans with his kids. "The blackberries are exploding right now," he says, "so we'll go pick some." ■ [SEE EDITORIAL P.5](#)

Ewen Callaway is a senior reporter for *Nature* in London.



THE LINGERING QUESTIONS

In the haze of incomplete data, scientists are divided over the risks and benefits of electronic cigarettes.

BY DANIEL CRESSEY

In many respects, the modern electronic cigarette is not so different from its leaf-and-paper predecessor. Take a drag from the mouthpiece and you get a genuine nicotine fix — albeit from a fluid wicked into the chamber of a battery-powered atomizer and vaporized by a heating element. Users exhale a half-convincing cloud of ‘smoke’, and many e-cigarettes even sport an LED at the tip that glows blue, green or classic red to better simulate the experience romanticized by countless writers and film-makers. The only things missing are the dozens of cancer-causing chemicals found in this digital wonder’s analogue forebears.

E-cigarettes — also known as personal vaporizers or electronic nicotine-delivery systems among other names — are perhaps the most disruptive devices that public-health researchers working on tobacco control have ever faced. To some, they promise to snuff out a behaviour responsible for around 100 million deaths in the twentieth century. Others fear that they could perpetuate the habit, and undo decades of work.

Now, a group once united against a common enemy is divided. “These devices have really polarized the tobacco-control community,” says Michael Siegel, a physician and tobacco researcher at Boston University School of Public Health in Massachusetts. “You now have two completely opposite extremes with almost no common ground between them.”

Evidence is in short supply on both sides. Even when studies do appear, they are often furiously debated. And it is not just researchers who are attempting to catch up with the products now pouring out of Chinese factories: conventional tobacco companies are pushing into the nascent industry, and regulators are scrambling to work out what to do.

Some countries, such as Singapore and Brazil, have banned the products entirely. The US Food and Drug Administration has proposed to bring them under its control alongside tobacco — but the path to regulation has been beset by lawsuits and delays. In May, the European

E-cigarettes are touted as a safe alternative to tobacco, but research has been inconclusive.

CHRISTOPHE ENA/AP PHOTO

Union finalized a major revision to the rules governing tobacco products in its member states. These include standards for e-cigarette products and restrictions on advertising, but the updated rules will take years to come into effect. On 26 August, the World Health Organization (WHO) released a report that recommended, among other things, to restrict the indoor use of e-cigarettes, to ban certain flavours and to confine sales to those who are 18 years and older. The report will be debated at a meeting in October to decide how the products are treated under the international Framework Convention on Tobacco Control, which commits governments to regulating tobacco and trying to reduce its impact on health.

The open questions include exactly what is in many commercially available products and what health effects they might have. But researchers are also concerned with whether e-cigarette users will give up conventional smoking, or simply become 'dual users'. Could e-cigarettes even act as a gateway, increasing tobacco use?

Siegel says that it is obvious what data and experiments are needed, but it is not guaranteed that anyone will agree about the results. "It's not clear to me that science is going to end this," he says.

MARKET CONFLAGRATION

Devices for taking the smoke out of smoking have been around for years, but most have failed to gain traction or, like prescription nicotine inhalers, are restricted in their use. A Chinese inventor named Hon Lik is widely credited for developing the modern e-cigarette, about a decade ago. The Shenzhen-based company he worked for, now called Ruyan, commercialized the invention and has been joined by scores of competitors.

According to a study¹ from the University of California, San Diego, there were 288 brands of e-cigarette available online in 2012, many with multiple products. By January 2014, there were 466, meaning that an average of more than 10 brands had been launched every month. Buyers have clearly been snapping them up: from a standing start a few years ago, the United Kingdom alone is now estimated to have more than 2 million users.

This explosive growth has blind-sided scientists and regulators alike. "I'm personally astounded by how quickly the market has grown," says Wilson Compton, deputy director of the US National Institute on Drug Addiction in Bethesda, Maryland.

Further complicating the picture is the remarkably fast evolution of the devices themselves. Early models that resemble cigarettes — 'cigalikes' — have been joined by customizable vaporizers costing hundreds of dollars and sporting everything from gold plating to software that lets users tweak how the devices operate.

In response, researchers have radically scaled up their efforts to provide regulators with guidance. E-cigarettes promise to drastically reduce the death toll from smoking — without depriving users of the nicotine they crave. (A phrase often quoted in tobacco-control circles is that people 'smoke for the nicotine but die from the smoke'.)

But on the central question — are e-cigarettes safe? — there are many uncertainties. Long-term consumption of nicotine divorced from tobacco is thought to be relatively safe for most people, barring pregnancy or certain rare conditions. But nicotine is not danger-free. There have already been overdoses from people drinking the liquid from e-cigarettes, or spilling it on their skin, where it is absorbed.

Also unknown are the long-term effects of regularly inhaling propylene glycol, the chemical that makes up most of the liquid vaporized in e-cigarettes. This organic molecule is used in scores of commercial applications ranging from food to plastics, and it has been shown to be safe to consume except at very high levels. Some evidence from the theatre — where it is used to create fogs and mists — suggests that it may irritate the respiratory

system, but there are no long-term data about the effects of inhalation.

Many e-cigarettes contain other chemicals added for flavouring, and little is known about these. There are also legitimate fears about quality standards for the products: toxic contaminants have been found, and in a very few cases batteries have exploded, leading to injury.

Around the world, researchers are now subjecting e-cigarettes to the same kinds of tests used to shed light on how conventional cigarettes damage human health. Some have found² genetic changes to human bronchial cells grown *in vitro* in a medium exposed to e-cigarette vapour (see *Nature* 508, 159; 2014). These looked similar to changes induced by conventional tobacco smoke. Another study found³ that e-cigarette use, like normal cigarette smoking, led to a reduction in exhaled nitric oxide, which could be a sign that e-cigarettes alter lung function. But this work is early and still inconclusive.

Those who are positive about the potential benefits of e-cigarettes say that although their safety clearly needs to be monitored and further investigated, there is simply no way they can be as dangerous as conventional cigarettes.

"The key comparison here is to smoking," says Lynne Dawkins, head of the drugs and addictive behaviours research group at the University of East London, UK. Dawkins says that the lower risks of e-cigarettes and the fact that many users believe they are an acceptable substitute for tobacco makes them generally a good thing. (Some of Dawkins's research has been funded by e-cigarette companies.)

SNUFF IT OUT

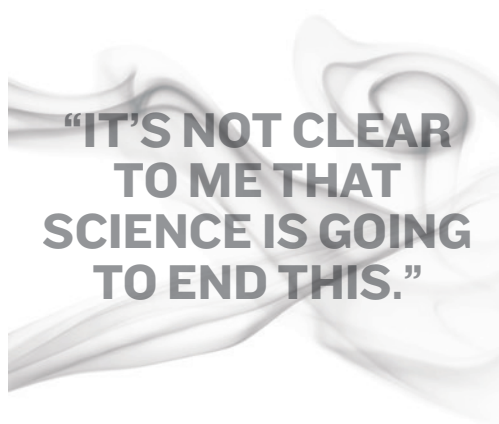
Dawkins and others are optimistic that beyond being a safer substitute, e-cigarettes could help people to stop smoking. But in many of the jurisdictions where they are taking off, e-cigarettes cannot be sold as smoking-cessation aids. In the United Kingdom, for example, that would require them to be licensed as medicine. The United States also bans direct claims about helping people to quit, but some brands circumvent this with testimonials from users or other implied messages about the devices' benefits.

So far there is a lot of anecdote but only a little hard evidence. One of the few randomized controlled trials on e-cigarettes comes from Christopher Bullen, who studies tobacco control at the University of Auckland in New Zealand. His study, published last year⁴, found that an early e-cigarette model was roughly as effective as nicotine patches at helping smokers to quit. But critics have cited weaknesses such as problems monitoring the actual use of devices and differences in how study participants obtained them. Participants may have had to go to more effort, for example, to obtain patches than to get e-cigarettes.

In the absence of further controlled trials, researchers have scoured the Internet for data and conducted surveys of smokers. Dawkins and her team have found⁵ that many people report using e-cigarettes for smoking cessation, and that these users had a longer 'time to first vape' in the morning than smokers had to their first cigarette, possibly suggesting reduced dependence on nicotine.

But opponents of e-cigarettes have their own ammunition. One study published this year⁶ followed 949 smokers reporting their habits online, and found that the e-cigarette users were no more likely to quit tobacco than other smokers. Dawkins and other e-cigarette defenders counter that because the devices may appeal to the smokers who are most heavily dependent on tobacco, results such as this do not actually shed light on the question (see *Nature* <http://doi.org/t3f>; 2014).

One problem with using e-cigarettes for smoking cessation may be that at the moment most are probably less effective at delivering nicotine than conventional smoked tobacco, says Peter Hajek, a tobacco



researcher at Queen Mary University of London. “I think they need about five years, if the regulators don’t kill them, to become as good as cigarettes in providing smokers with what they want.” This, he says, could ultimately render normal cigarettes obsolete.

But in many ways, those in favour of stricter controls on the devices are worried about giving up any ground in the fight against tobacco. As smoking becomes more difficult — for example, through restrictions on where smokers can light up — e-cigarettes may be used alongside conventional tobacco to maintain nicotine levels. Such dual use could undermine efforts to stop smoking entirely. And although dual users may consume fewer cigarettes than heavy smokers, which would reduce their risk of cancer to some extent, even very low levels of smoking seem to elevate risks of cardiovascular problems.

Those who worry that e-cigarettes will do more harm than good also fret that they could make tobacco socially acceptable again. With many developed nations implementing heavy restrictions on advertising, as well as high taxes and medical warnings, tobacco consumption has been massively stigmatized. Now e-cigarettes — which are in many cases unregulated — threaten to disturb this status quo.

One of the opponents’ greatest fears is that e-cigarettes will help to attract young people to tobacco. The US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, has found⁷ that in 2012, around 1.78 million adolescents in the United States had used e-cigarettes, and that a little less than 10% of those had never previously tried conventional cigarettes.

When those figures were released last year, CDC director Tom Frieden — who headed several anti-smoking initiatives in a previous role as New York City’s health commissioner — said that “the increased use of e-cigarettes by teens is deeply troubling”. And he warned: “Many teens who start with e-cigarettes may be condemned to struggling with a lifelong addiction to nicotine and conventional cigarettes.”

But teenagers often experiment, and it may be that this is all that these data show. Advocates of the devices say that if they were going to cause increases in smoking, then smoking rates would already be going up, given the number of people using e-cigarettes. This does not seem to have happened yet — in developed nations, smoking rates are generally decreasing.

YOUNG AND VULNERABLE

A contentious paper⁸ on this subject, and one that exemplifies the debate, comes from Stanton Glantz, director of the Center for Tobacco Control Research and Education at the University of California, San Francisco, who has spent years fighting tobacco and the industry that produces it.

In March, Glantz and his colleague Lauren Dutra analysed a survey of US adolescents and found that those who used e-cigarettes were more likely than others to smoke conventional cigarettes. They wrote that “in combination with the observations that e-cigarette users are heavier smokers and less likely to have stopped smoking cigarettes, these results suggest that e-cigarette use is aggravating rather than ameliorating the tobacco epidemic among youths”.

The paper drew strong criticisms for conflating correlation and causation. “These researchers are drawing conclusions that aren’t justified by the data,” says Siegel. Although there is clearly a correlation between heavy smoking and e-cigarettes, he says, it is not clear whether e-cigarettes are leading to smoking, or the other way around.

Glantz says that much of the ire directed at the paper is the result of the word ‘gateway’ being used in a press release, which he was unhappy with. He maintains that the data in the paper back the conclusion.

Overall, Glantz says, “properly regulated and available on prescription”, e-cigarettes might be a good thing, but they are currently increasing the number of children using nicotine, and promoting cigarette-smoking among children.

Parties on both sides of the debate had been petitioning the WHO even before it took its firm stance against the devices in August. In a 26 May letter to WHO head Margaret Chan, leading researchers



TYRONE SUI/REUTERS/CORBIS

Manufacture of e-cigarettes is booming in China.

including Dawkins, Bullen and Hajek argued that tough regulation would be counterproductive and would serve only to protect the conventional cigarette market. Harm-reduction approaches, they say, seem to have been “overlooked or even purposefully marginalized”.

Another group of equally eminent scientists — including Glantz — fired back in June, saying that there is insufficient evidence to show that e-cigarettes are useful for smoking cessation, that there is good evidence that they release toxic compounds, and that letting e-cigarettes go largely unregulated could once again allow tobacco companies the opportunity to influence policy.

Big tobacco is moving into the market with gusto. Leading US brand Blu — which Reynolds American, maker of Camel cigarettes, agreed to sell to rival Imperial Tobacco in July — has cornered about half of the US market by some estimates. Reynolds has kept hold of the popular VUSE brand. Altria, which is famous for the Marlboro cigarette brand, has its own MarkTen e-cigarette.

Jason Hughes, a tobacco researcher and head of the department of sociology at the University of Leicester, UK, notes that although e-cigarettes are often seen as something totally new, they may actually be just one more in a long line of attempts to make tobacco consumption more ‘civilized’, from chewing tobacco to cigarettes to cigarettes with filters. But they also represent a break point: although the nicotine in them is derived from plants, the users are now divorced from tobacco leaves completely.

Determining whether this break is truly a good thing becomes crucial when — despite continuous and graphic warnings of the risks of smoking — millions still put their lives at risk for a nicotine hit. Population studies to work out the true effects of this new technology are crucial, says Compton.

There is one thing that all researchers agree on: while they debate, e-cigarette use grows and grows. Whatever researchers think, says Compton, “The public is clearly voting with their feet.” ■

Daniel Cressey writes for Nature from London.

1. Zhu, S.-H. *et al. Tob. Control* **23**, iii3–iii9 (2014).
2. Park, S. J. *et al. Clin. Cancer Res.* **20**, B16 (2014).
3. Marini, S., Buonananno, G., Stabile, L. & Ficco, G. *Toxicol. Appl. Pharmacol.* **278**, 9–15 (2014).
4. Bullen, C. *et al. Lancet* **382**, 1629–1637 (2013).
5. Dawkins, L., Turner, J., Roberts, A. & Soar, K. *Addiction* **108**, 1115–1125 (2013).
6. Grana, R. A., Popova, L. & Ling, P. M. *JAMA Intern. Med.* **174**, 812–813 (2014).
7. Corey, C. *et al. Morb. Mortal. Wkly Rep.* **62**, 729–730 (2013).
8. Dutra, L. M. & Glantz, S. A. *JAMA Pediatr.* **168**, 610–617 (2014).

COMMENT

SATELLITES A call for all Earth observations to be open access **p.30**



BIOLOGY Lewis Wolpert's survey of sex differences, reviewed **p.32**

ENERGY Social sciences and humanities take their seats at the table **p.33**

OBITUARY Yoshiki Sasai, stem-cell pioneer, remembered **p.34**

ILLUSTRATION BY CHRIS RYAN/NATURE



Where is the brain in the Human Brain Project?

Europe's €1-billion science and technology project needs to clarify its goals and establish transparent governance, say **Yves Frégnac** and **Gilles Laurent**.

Launched in October 2013, the Human Brain Project (HBP) was sold by charismatic neurobiologist Henry Markram as a bold new path towards understanding the brain, treating neurological diseases and building information technology. It is one of two 'flagship' proposals funded by the European Commission's Future and Emerging Technologies programme (see go.nature.com/icotmi). Selected after a multiyear competition, the project seemed like an exciting opportunity to bring together neuroscience and IT to generate practical applications for health and medicine (see go.nature.com/2eocv8).

Contrary to public assumptions that the HBP would generate knowledge about how the brain works, the project is turning into an expensive database-management project with a hunt for new computing architectures. In recent months, the HBP executive board revealed plans to drastically reduce its experimental and cognitive neuroscience arm, provoking wrath in the European neuroscience community.

The crisis culminated with an open letter from neuroscientists (including one of us, G.L.) to the European Commission on 7 July 2014 (see www.neurofuture.eu), which has now gathered more than 750 signatures.

Many signatories are scientists in experimental and theoretical fields, and the list includes former HBP participants. The letter incorporates a pledge of non-participation in a planned call for 'partnering projects' that must raise about half of the HBP's total funding. This pledge could seriously lower the quality of the project's final output and leave the planned databases empty.

With the initial funding, or 'ramp-up', phase now in full swing, the European Commission is currently evaluating the HBP directors' plan for the larger second part of the project. This offers an opportunity to introduce reforms and reconciliation. ►

► Here, we offer our analysis of how the HBP project strayed off course and how it might be steered back.

THE ROOTS OF CRISIS

The HBP blends two styles. One comes from a history of successful interdisciplinary collaborations in the European Union in brain- and neuron-inspired computation¹. The second originates from a computational research programme, the Blue Brain Project², initiated by Markram in 2005 (see 'Brain activity'). This collaboration between the Swiss Federal Institute of Technology in Lausanne (EPFL) and the IBM computing corporation aimed to build large-scale 'bottom up' numerical simulations of a rat's neocortical column, a set of about 100,000 neurons considered to be a functional unit within the brain.

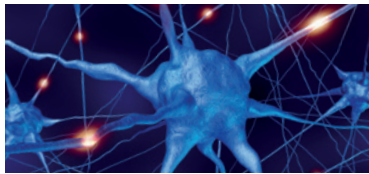
The crisis results mainly from ambiguities concerning the place of neuroscience in the HBP. From the beginning, neuroscientists pointed out that large-scale simulations make little sense unless constrained by data, and used to test precise hypotheses. In fact, we lack, among other resources, a detailed 'connectome', a map of connections between neurons within and across brain areas³ that could guide simulations. There is no unified format for building functional databases or for annotating data sets that encompass data collected under varying conditions. Most importantly, there are no formulated biological hypotheses for these simulations to test⁴.

Many scientists also feared that the HBP would siphon funds from fundamental research. The European Commission's investment in a large 'brain project' would influence what other research areas it chooses to fund. Nonetheless, such an opportunity seemed unlikely to arise again, and neuroscientists (ourselves included) joined up, even if they did not agree with all aspects of the HBP proposal or with certain promises used to sell it. We put our faith in open and interdisciplinary collaboration, trusting that intellectual and operational details would take shape gradually and collectively.

Preparation of the HBP flagship document (in a pre-project phase funded by the European Commission) lasted several months and took place at the EPFL. Selection criteria for flagship proposals included scientific vision, societal impact and the size of the scientific community involved. We trust that the selection of the HBP by the commission rested on its focus on the relationship between brain structure and function, and on the project's interdisciplinary approach. More than 240 labs were initially pegged to participate. Official descriptions expressed hopes that the project would "gain fundamental insights into what it means to be human, develop new treatments for brain diseases and build revolutionary new information and communications technologies". The exploration of the

BRAIN ACTIVITY

Timeline of the Human Brain Project.



2005 The European Union starts funding research merging computing architecture and neuroscience.

MAY 2005 Neurobiologist Henry Markram launches the Blue Brain Project, led by IBM and the Swiss Federal Institute of Technology in Lausanne (EPFL).

NOVEMBER 2008 US defence department launches programme on brain-inspired electronic systems.

JULY 2009 Markram publicizes his vision to build the brain in a supercomputer.

JULY 2010 The European Commission calls for proposals for 10-year, €1-billion 'flagship' interdisciplinary technology projects.

APRIL 2012 European neuroscientists sign on to a proposal for a brain project.

JANUARY 2013 The Human Brain Project (HBP) is selected as a winning flagship proposal, along with a project on graphene.

APRIL 2013 US President Barack Obama announces the BRAIN initiative.

OCTOBER 2013 The HBP launches, coordinated by Markram at the EPFL.

APRIL 2014 The HBP governance announces plans for the project's larger second phase, radically reducing the role of experimental neuroscience.

JULY 2014 European researchers write an open letter (now with about 750 signatories) to the European Commission, decrying shifting goals and lack of transparent leadership.

principles of neural computation through animal studies was considered integral.

Flagship initiatives differ from other European Commission projects because unusually large administrative, scientific and strategic leadership responsibilities fall onto few

principal investigators. As the HBP prepares to move into its second phase, the project's executive board has revised its objectives. As well as decreasing the emphasis on experimental neuroscience, it has eliminated non-human-primate research, and restricted the focus of experimentation mainly to human imaging and 'atlases'. These are typically static catalogues of gene expression, neuronal couplings, cell types and other measurements across brain structures, but without experiments to assess function. The board also announced plans to dissolve the cognitive-neuroscience sub-programme, leading to the resignation of that project's 18 principal investigators, including its director.

Neuroscience in the HBP is now limited mainly to simulations and to building a massive infrastructure to process mostly existing data. The revised plan advances a concept in which *in silico* experimentation becomes a "foundational methodology for understanding the brain"⁵. Numerical simulations and 'big data'⁶ are essential in modern science, but they do not alone yield understanding. Building a massive database to feed simulations without corrective loops between hypotheses and experimental tests seems, at best, a waste of time and money. The HBP's goals now look like a costly expansion of the Blue Brain Project, without any further evidence that it can produce fundamental insights.

BRAIN WRECK

Neuroscientists who initially supported the HBP feel that they have been taken advantage of. The organizers attracted well-funded neuroscience labs for credibility and, ultimately, for their data. Now those labs are being edged out.

The changes to the HBP are not only disingenuous, they are self-defeating. About €430 million (US\$570 million) of the European Commission funding goes to the HBP's 'core team'. The remainder of the €1-billion budget depends mostly on scientists throughout Europe raising partnering funds from sources such as regional governments, and then being selected by the HBP management on the HBP's terms. Why would people want to join the project under such conditions?

Since problems surfaced, the HBP executive board of directors and administrators at the EPFL, the coordinating institution, have been deflecting rather than addressing criticism (see go.nature.com/nenowj). Supporters argue that the Human Genome Project was also initially criticized by the biological community and eventually proved its detractors wrong. But the genome project was different: its goals were well defined, and the associated challenges, mostly technological, were well posed. These descriptions do not apply to the HBP.

In July, HBP co-executive director Richard Frackowiak wrote⁷ that the project is "a CERN

for the brain", equating the HBP with Europe's particle-physics laboratory near Geneva, Switzerland. But CERN is a cooperative structure built around large, shared instruments, designed to collect experimental data and test carefully constructed hypotheses. This, again, does not apply to the HBP.

Public grants of even a few tens of thousands of euros are typically subject to rigid oversight. Science projects of the HBP's magnitude and complexity are typically handled by agencies (such as the European Space Agency or CERN) with existing infrastructure and managerial experience. Why, then, did the commission grant nearly all of the design, management and decision power of a huge European public project to one academic institution and three principal investigators, two of whom are faculty members at the institution? The other flagship project, which focuses on graphene and has generated little controversy, opted for a distributed and transparent governance that includes all sub-project leaders on an executive board.

The tight schedule imposed to prepare the second phase of the HBP, compared with the several years taken to develop the proposal, has also exacerbated problems. How can a single group coordinate tens or hundreds of labs and sub-projects in such a short time, without tried-and-tested structures that facilitate the task and maintain trust?

Given the sizeable fraction of the HBP's core budget devoted to administration, we cannot feel assured that the HBP is managed cost-effectively or adapted for research operations across Europe. We are concerned that not all internal partners have read or had access to the 180-page draft of the next partnership agreement, and few seem aware of plans for a private, Swiss-operated foundation charged with exploiting commercial opportunities that emerge from the HBP.

WHERE DO WE GO FROM HERE?

We see three possible routes forward. All require transparent discussion of the project's goals and profound changes in governance and oversight.

One option would be to explicitly eliminate neuroscience from the HBP. This solution would annihilate the enormous efforts of many HBP partners (neuroscientists and clinicians), and redirect this flagship towards purely technological objectives. Finding a new name would be a small price to pay.

A second option would be to split the technology and neuroscience sections of the HBP. This would group all IT components of the HBP together and create of a new, independent entity to fund collaborative neuroscience to decipher brain function.

"We cannot feel assured that the HBP is managed cost-effectively."

The third option is to attempt to put the HBP back on track. This road is probably the most challenging. It would require re-establishing the HBP's experimental-neuroscience component with real funding avenues for biological and hybrid projects combining theory and experiment, under funding, review, award and administration conditions acceptable to the science community. Radical action would be required to revive trust and enthusiasm in an exceptional collective effort to address an enormously exciting challenge. Reunification might even be possible with major theoretical neuroscience institutes — the Gatsby Computational Neuroscience Unit in London, the Bernstein Centers in Germany, the École Normale Supérieure in Paris and the Edmond & Lily Safra Center for Brain Sciences in Jerusalem — at present largely hostile to the HBP. A Europe-wide committee of scientists should be established to organize neuroscience funding in Europe and work with funders to institute borderless, collaborative, curiosity-driven, interdisciplinary, peer-reviewed science. Issues concerning the partnership programme need to be addressed: notably the management and exploitation of data obtained through other funding sources.

The irony of this episode is that the HBP, by giving the impression of exceptional investment from the European Union in brain research, spurred the creation of competing, well-funded neuroscience initiatives in the United States (which focuses on techniques development⁸) and in China (which focuses on brain disease⁹). Now it is not at all clear that Europe has invested in brain science through the HBP. ■

Yves Frégnac is a CNRS research director and head of the CNRS Unit of Neuroscience, Information and Complexity in Gif-sur-Yvette, France. **Gilles Laurent** is director at the Max Planck Institute for Brain Research in Frankfurt, Germany.
e-mails: fregnac@unic.cnrs-gif.fr; gilles.laurent@brain.mpg.de

1. European Commission. *Future & Emerging Technologies (FET): FP7 Projects Compendium 2007–2013* (European Union, 2013).
2. Markram, H. *Nature Rev. Neurosci.* **7**, 153–160 (2006).
3. Seung, S. *Connectome: How The Brain's Wiring Makes Us Who We Are* (Houghton Mifflin Harcourt, 2012).
4. Frégnac, Y. *Cell* **155**, 265–266 (2013).
5. Markram, H. et al. *Procedia Comp. Sci.* **7**, 39–42 (2011).
6. Kandel, E. R., Markram, H., Matthews, P. M., Yuste, R. & Koch, C. *Nature Rev. Neurosci.* **14**, 659–664 (2013).
7. Frackowiak, R. 'Defending the grand vision of the Human Brain Project' *New Scientist* (16 July 2014).
8. Brain Research through Advancing Innovative Neurotechnologies (BRAIN) working group. *BRAIN 2025: A Scientific Vision* (National Institutes of Health, 2014).
9. Poo, M.-M. *Natl Sci. Rev.* **1**, 12–14 (2014).

For a list of further reading on this topic, see go.nature.com/7jg7zc.



NASA

The Landsat 8 satellite launched in February 2013 (artist's impression).

Make Earth observations open access

Freely available satellite imagery will improve science and environmental-monitoring products, say **Michael A. Wulder** and **Nicholas C. Coops**.

Changes in land cover affect the global climate by absorbing and reflecting solar radiation, and by altering fluxes of heat, water vapour, carbon dioxide and other trace gases. Detailed assessments — regional, global, daily and seasonal — of land use and land cover are needed to monitor biodiversity loss and ecosystem dynamics and to aid in reducing emissions from deforestation and forest degradation¹.

Satellite imagery is the best source of such data, especially over large areas. Observations need to be extensive, regular and consistent to establish baselines and trends. But today, most satellite observations have limited coverage and compatibility, because they are controlled by the diverse objectives of national space programmes. In many cases, satellite data are restricted or charged for.

A new era of open-access satellite data has arrived. In 2008, the US Geological Survey (USGS) released for free to the public its Landsat archive, which dates back to the 1970s and is the world's largest collection of Earth imagery². Greater computing power

is also enabling scientists to manipulate big data representing larger areas and with greater sophistication, to produce multi-billion-pixel composite maps of land cover and change across regions, continents and the globe. Monitoring land-cover change in near-real time is now a reality.

Obstacles remain. Data coverage in the Landsat archive is variable, in both space and time (see 'Global coverage'). And few people have enough computing power and bandwidth to download and manipulate the data. Decision-makers remain largely unaware of the vastly improved opportunities for environmental monitoring offered by the latest methods, and so are not yet using such data to their full potential.

Scientists and policy-makers can support the shift to open-access satellite data, and coordinate efforts to deliver the detailed global monitoring required by international climate change and emissions-reduction programmes. Further, governments should open up their national satellite image archives and integrate compatible data to

fill gaps. And satellite imagery from future missions should be freely accessible to all to promote innovation and use.

Landsat is the longest-running civilian Earth-observing programme. The United States launched the first satellite in 1972, when public interest in space missions was high. In 1982, Landsat 4 began to deliver more-detailed imagery (with a spatial resolution of 30 metres, now considered a benchmark for historical analyses) and spectral channels at visible, near- and shortwave-infrared wavelengths, to track the unique signatures in reflected light of different types of vegetation.

The latest in the unbroken series of Landsat satellites is Landsat 8, launched in 2013, which measures Earth's surface with a resolution of 15–100 metres from visible to infrared wavelengths. Landsat data were archived from the outset, but early observing strategies and limitations to on-satellite storage and downlink capacity resulted in uneven global coverage.

Access policies have changed over the years. In the 1980s, Landsat observations

were commercialized. Fewer images were acquired and prices rose, reaching thousands of dollars each³ and curtailing use. When Landsat 7 was launched in 1999, it collected images across the globe more systematically. Redistribution of purchased images became permissible, and users formed consortia to bulk buy and share raw images and products.

Usage rocketed in 2008, when Landsat made its images free. More than a million images were downloaded in the first year, compared with a previous annual high of about 25,000 images sold. More than 20 million images have been downloaded since the archive opened and the rate continues to increase.

DATA REVOLUTION

The mass of satellite data is fuelling a revolution in processing and analysis techniques. Cloud cover, the bane of optical remote sensing, can be overcome by building composite images. Over a period, clean pixels are selected, avoiding the requirement that entire images be clear. This approach is transforming continental and global-scale monitoring⁴.

Cycles of disturbance and growth can also be captured by tracing clean pixels through time. Landscape changes after a fire can be checked against expectations for vegetation recovery, for instance. And logical sequences — such as transitions from young forest to old — can be included with greater confidence in models and projections.

Advances in visualization and distributed cloud computing are also changing the game. For example, the powerful parallel processing facilities of the NASA Earth Exchange⁵ and the Google Earth Engine can handle large volumes of geospatial and remotely sensed data. Such central facilities allow users to bring algorithms to the large data sets while minimizing duplication of storage and processing efforts. Developing regions such as Africa, which lack computing power and infrastructure⁶, can also benefit from such shared platforms if they have the partnerships and bandwidth to access them.

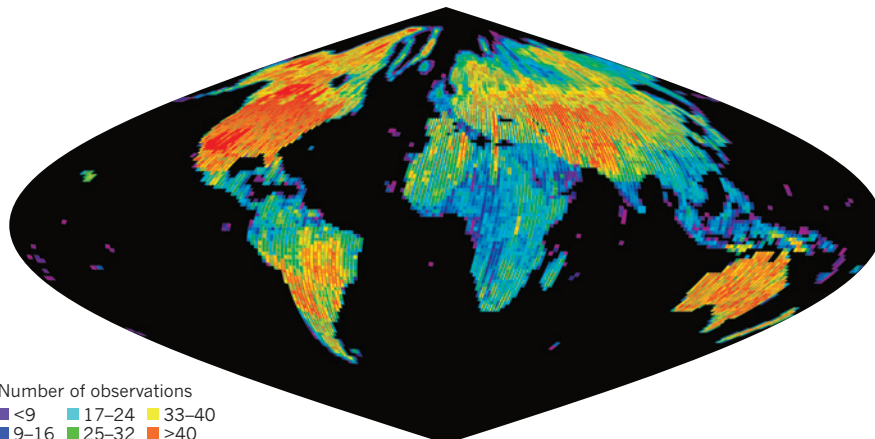
When the archive was opened, there were more Landsat images outside it than in it. Many images were retained by the global network of receiving stations. An effort to consolidate these has added more than 3 million images to the repository since 2010; agreements are in place for a further 2 million to be ingested.

Future coverage should be much better. Landsat 8 can collect more than 700 images per day — 14 times as much as in the 1980s. In April 2015, it should be joined by the first European Sentinel-2 satellite, which will map Earth's land area every 10 days. A second Sentinel-2 satellite is scheduled to be launched two years later. The three satellites together will image global land area twice a week.

Sentinel-2 also has a free and open data

GLOBAL COVERAGE

More than 120,000 Landsat images taken from December 2009 to November 2010 cover most of Earth's surface. On average, each 185 × 185-kilometre frame is imaged 21 times per year. But, irregular data receiving and uploads to the archive have resulted in patchy coverage. Current and future satellites will map the land more evenly.



Number of observations
 ■ <9 ■ 17–24 ■ 33–40
 ■ 9–16 ■ 25–32 ■ >40

policy, although at present the data will be publicly available for only four months after collection. This strategy would limit access to wide timespans of data and could result in incomplete repositories and multiple versions of data with different calibrations.

INTERNATIONAL STRATEGY

Governments and the remote-sensing community should now seize the opportunity to develop a unified strategy for land monitoring.

First, as much historical imagery as possible should be deposited in the Landsat archive or an equivalent open repository. Leading Earth-observation nations⁷ whose satellites have sensors that complement Landsat should explore opportunities to add their data to the global pool. Second, the Earth-observing community and govern-

“Landsat 8 can collect more than 700 images per day — 14 times as much as in the 1980s”

ments should commit to making future satellite programmes open access whenever possible. We urge the European Space Agency to consider ways to widen access to Sentinel-2's images.

Alternative delivery mechanisms, such as a third-party archive and distribution site, could make data available to a broader array of users, allowing the intended science benefits to be reached.

Earth-observation projects need to become more like meteorological programmes, with standard imaging specifications and internationally agreed complementary satellite launches to ensure continuous coverage. The costs and difficulties in integrating national programmes for international benefit are significant. The benefits of open access demonstrated by the Landsat programme justify and encourage these efforts. We urge the Committee on Earth Observation Satellites interagency forum to continue facilitating

communication between space agencies, commercial entities and scientific and operational data users towards establishing a more unified programme.

The remote-sensing community must advocate the development and maintenance of data archives and innovative processing methods. Best-practice approaches and standards can be developed by the Group on Earth Observations, a voluntary partnership between governments and international organizations to promote global collaboration around Earth observations, and offshoots such as the Global Forest Observation Initiative. To facilitate the use of data from differing sensors, scientists and data distribution centres should offer users calibrated data in widely compatible, analysis-ready formats.

Finally, researchers, policy-makers, non-governmental organizations and land managers need to use and promote more widely the capacity of the satellite archives for capturing and characterizing past and present changes in land cover and land use. ■

Michael A. Wulder is a senior research scientist at the Canadian Forest Service, Natural Resources Canada, Victoria, Canada. **Nicholas C. Coops** is professor of remote sensing at the Faculty of Forest Resources Management, University of British Columbia, Vancouver, Canada. e-mail: mike.wulder@nrcan-rncan.gc.ca

1. Phelps, J., Webb, E. L. & Adams, W. M. *Nature Clim. Change* **2**, 497–503 (2012).
2. Woodcock, C. E. *et al. Science* **320**, 1011 (2008).
3. Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R. & Woodcock, C. E. *Remote Sens. Environ.* **122**, 2–10 (2012).
4. Hansen, M. C. & Loveland, T. R. *Remote Sens. Environ.* **122**, 66–74 (2012).
5. Nemani, R., Votava, P., Michaelis, A., Melton, F. & Milesi, C. *EOS* **92**, 109–110 (2011).
6. Roy, D. P., Ju, J., Mbaw, C., Frost, P. & Loveland, T. *Remote Sens. Lett.* **1**, 111–117 (2010).
7. Belward, A. S. & Skoien, J. O. *ISPRS J. Photogramm. Remote Sens.* <http://dx.doi.org/10.1016/j.isprsjprs.2014.03.009> (2014).

Splitting the sexes

Virginia Valian ponders a study on biology, evolution and gender differences in humans.

Depending on where you buy it, this book's title is either *Why Can't a Woman Be More Like a Man?* (United Kingdom), or *Why Can't a Man Be More Like a Woman?* (United States). In my view, both titles suggest the retort: each can be.

Lewis Wolpert, a developmental biologist, surveys evidence for historical, neural, hormonal, emotional, cognitive, linguistic and health differences and similarities between men and women in an attempt to shed light on their consequences. That is a lot of ground to cover in 224 pages — a bit like *The Reduced Shakespeare Company's* 80-minute *The Complete Works of William Shakespeare (Abridged)*. In the 22-page chapter on sex and reproduction, for instance, Wolpert reviews sexual activity and arousal, choice of sexual partners, cross-cultural differences in mating, sexual orientation and transsexuality. The ten-page chapter on sex discrimination takes us from Herodotus to women's underrepresentation in maths and science. Nuance is inevitably lost, as is an organizing principle.

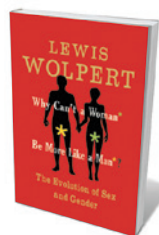
One can learn a lot from Wolpert about biological sex differences, but the book never overcomes one difficulty. The very data that Wolpert presents — such as the minimal differences between men and women in skills, cognition and language — undermine his thesis that brain and hormonal differences are major contributors to disparities in professional success.

Wolpert notes, for instance, that few of the purported differences between male and female brains, such as the extent of grey and white matter, seem linked to differences in male and female behaviour. Despite such statements, and a heavy use of conditionals such as 'may' and 'might', Wolpert still expresses views that are at best over simplified and at worst cartoonish. For example, he repeats the observation that "people everywhere understand sex as 'something females have that males want'".

In his chapter on skills, Wolpert notes that one of the few robust sex differences in cognition is the superior ability of males, on average, to mentally rotate three-dimensional (3D) images. This is linked to hormones. But behavioural effects of hormonal differences are not straightforward, despite thorough documentation of male and female hormones and investigations of children with congenital adrenal hyperplasia who have been exposed to atypically high

amounts of testosterone in the womb. In girls, one consequence is a superior ability to rotate 3D images, but boys show reduced ability. Although that supports the likelihood that spatial differences in unaffected males and females are mediated by hormones, they are not dictated by them — training can eliminate the difference.

To clarify the paucity of cognitive differences, Wolpert introduces a helpful perspective from the developmental psychologist Uta Frith: equivalent behaviours need not require the same neural mechanism. (Adding and subtracting, after all, can be accomplished by calculators with different mechanisms.) In a similar vein, neuroendocrinologist Geert De Vries and others have conjectured that neural and hormonal sex differences can interact to eliminate or



Why Can't a Woman Be More Like a Man?
LEWIS WOLPERT
Faber and Faber:
2014.

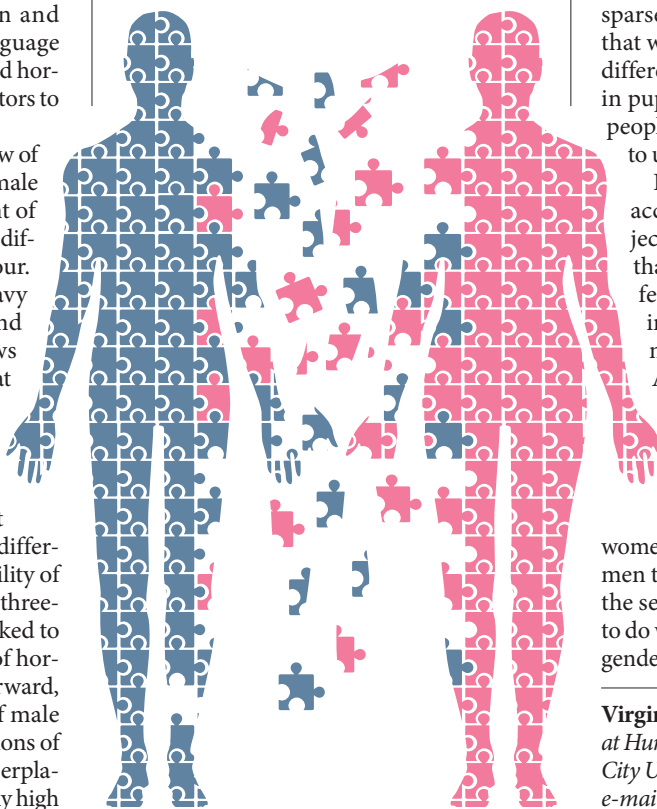
create behavioural differences. There is a fascinating question here about the relation between brain, hormones and behaviour, which would have produced a richer book had Wolpert pursued it.

With respect to women's maths abilities and under-representation in science, Wolpert offers a grab-bag of facts with little interpretation. He also fails to consider the implications of the social and political factors inherent in transnational differences, such as China having a much higher percentage of women in engineering than the United States. Instead, and despite mixed evidence, Wolpert proposes that empathy "marks a fundamental difference" between the sexes and suggests that as a reason for men's dominance in science.

Empathy, however, is a complex state involving both cognitive components such as the ability to recognize others' emotional states, and emotional components such as the ability to share the emotions involved in those states. Women report greater feelings of empathy than men do, but that may reveal little, because both sexes know that women are expected to be more empathetic. In well-controlled studies, sex differences are sparse. Some research finds, for example, that whereas men and women self-report differently, both show equivalent increases in pupil dilation when viewing images of people being hurt deliberately as opposed to unintentionally.

In his conclusion, Wolpert tries to account for "the almost universal subjection of women by men", speculating that along with factors such as size differences, pregnancy and breastfeeding, "empathy possibly makes women more willing to do what men request". Although it is qualified, that last statement comes uncomfortably close to painting women as the unprotesting victims of selfish men. By seeing sex differences in empathy as close to immutable, Wolpert relegates women to suboptimal professional lives and men to suboptimal parenting. What keeps the sexes from being more alike has more to do with social structure and beliefs about gender differences than with bare biology. ■

Virginia Valian is a distinguished professor at Hunter College in New York City and the City University of New York Graduate Center. e-mail: virginia.valian@hunter.cuny.edu



SHUTTERSTOCK

Correspondence

Energy: social fixes are on the up

Already governments, environmental agencies, non-governmental organizations and businesses are on the way to meeting Benjamin Sovacool's call for energy research to be more socially oriented (*Nature* **511**, 529–530; 2014). The annual Behavior, Energy and Climate Change Conference, this year to be held in December, reviews their achievements and ongoing research activities.

Examples of social energy practices include smartphone apps that allow remote control of domestic thermostats and can even track carbon emissions. The US Energy Star programme and the Leadership in Energy and Environmental Design rating systems ensure that buildings are energy-efficient and environmentally sound. Other energy-efficiency initiatives include Japan's Top Runner and the UK Energy Saving Trust.

The US company Opower, which serves 32 million households and businesses, enables consumers to compare their energy usage. Car manufacturers, with help from researchers at the University of California, Davis, are studying drivers' behaviour with a view to creating a display that indicates how to cut fuel consumption. And US start-up companies such as Zipcar, Uber and Lyft are changing consumer relationships with their vehicles by enabling car sharing.

Developing countries are also on the front line of new business models for clean and sustainable energy. The 'en.lighten' initiative, led by the Global Environment Facility, the United Nations Environment Program and industry partners, is accelerating conversion to energy-efficient lighting. In India, the think tank TERI (The Energy and Resources Institute) is leading efforts to provide solar lighting for poor communities and tackling gender issues that might prevent it from

being widely accepted. And in Africa, mobile-phone payment systems are making it easier for kerosene consumers to switch to solar energy.

David E. Rodgers, Dustin S. Schinn *The Global Environment Facility, Washington DC, USA.*
dschinn@thegef.org

Energy: we need all hands on deck

A more interdisciplinary approach will advance our understanding of energy issues (see also B. K. Sovacool *Nature* **511**, 529–530; 2014).

In considering how to improve energy efficiency using a particular technology, for example, an engineering-only analysis would seek to optimize technical efficiency, whereas an engineering–economic analysis would aim to make it more cost-effective. But an interdisciplinary analysis would take in wider considerations. For example, convenience to consumers and the attributes of organizations promoting the technology can affect adoption rates by a factor of ten when financial incentives are comparable (P. C. Stern *Energy Res. Soc. Sci.* **1**, 41–48; 2014).

Psychologist Kurt Lewin remarked that “there is nothing so practical as a good theory” (*Field Theory in Social Science* Harper, 1951). I suggest that nothing advances theory better than tackling a practical problem by integrating different perspectives.

Paul C. Stern *National Research Council, Washington DC, USA.*
psstern@nas.edu

Energy: humanities frame the aims

Social scientists offer more than just information for producing a best fit between energy technologies and social circumstances (B. K. Sovacool *Nature* **511**, 529–530;

2014). By eliciting different social framings of energy systems and technologies, they help to determine points of commonality and disagreement about the sort of Earth we want to inhabit.

The humanities are concerned with profound questions about justice, fairness, responsibility and care. Such questions can no longer be regarded as external to the design of energy devices and systems because they are central to the entire social and technical fabric of modern society (see, for example, N. Castree *et al. Nature Clim. Change* **4**, 763–768; 2014).
Noel Castree *University of Wollongong, Australia; and University of Manchester, UK.*
ncastree@uow.edu.au

Better monitoring of fish in dam projects

Inadequate ecological monitoring before and after removal of dams in the northwestern United States is hampering evaluation of the success of dozens of these projects (see *Nature* **511**, 521–523; 2014).

A primary purpose of dam removal in the region has been to revive populations of endangered migrating salmonid fish by restoring natural river functions. Populations are being monitored after demolition of the 38-metre Condit dam and 64-metre and 33-metre Elwha dams, but these are exceptions.

Using fish ladders, salmonids were able to ascend the Marmot and Hemlock dams you mention and move into the upper river. But to gauge the ecological impact of such dam removals, it is necessary to have baseline data on fish passage from before the structure was destroyed. In most other cases, these data are not available, thwarting future assessments.

Long-term ecological evaluation and monitoring must be built into dam-removal projects. Otherwise, millions of

dollars will continue to be spent with no evidence that the desired outcome is being achieved.

Peter Brewitt, Karen D. Holl *University of California, Santa Cruz, USA.*
pbrewitt@ucsc.edu

Seals collect more Southern Ocean data

Marine mammals equipped with recording devices can help to overcome the practical challenges of collecting year-round physical data in the Southern Ocean (see M. C. Kennicutt II *et al. Nature* **512**, 23–25; 2014).

Data collection using satellite remote sensing or surface research vessels, for instance, is limited to the short summer period for access to high latitudes or to the northerly margins of winter sea ice.

By contrast, small oceanographic sensors worn by seals can provide high-resolution oceanographic profiles from Antarctica throughout the winter (see, for example, F. Roquet *et al. Scientific Data*, in the press). The data can be used to study and map ocean fronts and to monitor the annual formation and growth rate of sea ice (J.-B. Charrassin *et al. Proc. Natl Acad. Sci. USA* **105**, 11634–11639; 2008).

Data derived using seals can help us to predict rapid changes and their effects in the Southern Ocean. This would give the global community time to develop adaptation policies.

Clive R. McMahon *Sydney Institute of Marine Science, New South Wales, Australia.*
Robert Harcourt *Macquarie University, Sydney, Australia.*
clive.mcmahon@utas.edu.au

CONTRIBUTIONS

Correspondence may be submitted to correspondence@nature.com after consulting the guidelines at <http://go.nature.com/cmchno>.

Yoshiki Sasai

(1962–2014)

Stem-cell biologist who decoded signals in embryos.

How does a fertilized egg — a single cell — produce the myriad specialized cells that assemble into three-dimensional tissues and functioning organs? This fundamental secret has captivated generations of embryologists. Using creative culture conditions and unmatched insight into the action of biological signals that govern tissue development, Yoshiki Sasai found ways to mimic these mysterious processes in the laboratory. He leaves a legacy of remarkable discoveries that expand the frontiers of stem-cell research and tissue regeneration.

Yoshiki struck me as a happy scientist. He spoke softly and with a unique smile as he described Japanese traditions or revealed his astonishing findings. My first experience of his pride and passion came during a long walk on Japan's Rokko Mountain in 2008. We looked over the splendour of Kobe and the reclaimed island where his institute, the RIKEN Center for Developmental Biology, lay. The sight brought that smile to his face. As he taught me the art of ninja walking, recounted the local thousand-year-old methods used to make sake, and dissected the etiquette of Japanese bathing, he explained some of his incredible scientific discoveries.

Sasai was a master at deciphering the code by which cells learn their place in a developing embryo. He tested this code on cultures of embryonic stem (ES) cells derived from early-stage embryos, and instructed these unspecialized cells into becoming specific types of neuron. He made nerve cells found in different parts of the forebrain by varying concentrations of morphogens, molecules that guide the patterning of tissues. Sasai's laboratory developed the first methods to make inhibitory interneurons, which hold promise for brain repair. He also discovered a simple set of cues to make the brain's hypothalamic neurons and pituitary cells, important for several bodily functions.

Sasai instructed both mouse and human ES cells to become cortical neurons, the type that expanded markedly during the evolution of the human brain. Remarkably, his ES-cell-derived cortical cells assembled into floating, three-dimensional organoids similar to the embryonic organ from

which the brain's cerebral cortex arises. By carefully adding growth factors, he coaxed the organoids into forming the front or the back of the cortex. These advances provide powerful tools for understanding healthy and aberrant human brain development. The methods that he developed may one day be used to make replacement cells to treat currently incurable diseases.



Sasai amazed the stem-cell community with time-lapse videos showing human and mouse ES cells turning into optic vesicles, the precursors of eyes. He demonstrated that these go on to form an optic cup and all the major cell types in the retina. Sasai's laboratory inferred the basic principles of how eyes form from self-directed cell assembly, gaining insights into a process that is difficult to monitor and manipulate in developing animals. When I visited his laboratory in February, Sasai shared recent work indicating how these optic cups acquired polarity, in which individual cells take on different properties according to their position in a larger structure. This trait is crucial for neural connectivity and vision.

Sasai's work deriving cortical, visual and pituitary tissues demonstrated that complex, highly organized structures could be made from isolated human ES cells. The work broke new ground in understanding tissue and organ formation, modeling disease and advancing regenerative

medicine. He thought that it could be pushed even further, thinking of ways to generate multiple brain structures *in vitro* that could become interconnected.

Sasai obtained his MD in 1986 and his PhD in 1993 from Kyoto University, where he trained as a molecular biologist in the laboratory of neuroscientist Shigetada Nakanishi. He then went to the University of California, Los Angeles, and worked with embryologist Edward De Robertis to discover chordin, one of the key early inducers in the formation of the nervous system. In 1996, Sasai moved back to Kyoto University, where he began his extraordinary work mimicking embryonic development in a dish. He moved to RIKEN in 2000.

Given the pride that Yoshiki felt for the work he was doing in his home country, his enjoyment of and devotion to good science, his family, and life in general, I was shocked to learn that on 5 August my admired colleague and dear friend had taken his own life. Clearly, his mental state took a very heavy blow from the onslaught of media attention and the months of allegations surrounding two *Nature* papers published in January from the laboratory of Haruko Obokata, on the generation of stimulus-triggered acquisition of pluripotency (STAP). Laboratories around the world were unable to reproduce the findings, and after serious problems were found with the studies, the papers were retracted in July. Although the work was not a direct product of his laboratory, Sasai was a co-author of the papers. In June, he wrote to me expressing his deep concerns about the worsening STAP situation in Japan.

Yoshiki's career was full of insights that prompted the unique smile that we will greatly miss. I hope that his realized and unrealized ideas inspire the next generations of scientists to follow his solid legacy of research. It is a true tragedy that the scientific community is left without this forward thinker. ■

Arturo Alvarez-Buylla is professor of neurological surgery at the University of California, San Francisco, USA. He collaborated with Yoshiki Sasai between 2008 and 2014.
e-mail: abuylla@stemcell.ucsf.edu

COURTESY OF ARTURO ALVAREZ-BUYLLA

How plumes help to break plates

Computer models show how hot material that rises from Earth's interior is affected by plate tectonics, producing unexpected irregularities in Earth's topography and assisting in the break-up of continental plates. [SEE LETTER P.85](#)

SUSANNE BUITER

High mountains and deep valleys are eye-catching features of our continents. Such differences in surface elevation are the result of the converging and diverging motions of the tectonic plates that form Earth's strong outer layer — the lithosphere. Superimposed on this landscape is low topography over wide regions that is generated by slow movement in Earth's interior. How do these two processes interact to shape the ground beneath our feet? On page 85 of this issue, Burov and Gerya¹ demonstrate how plate tectonics turns symmetric deep mantle flows into irregular surface topography, thereby touching on the debate over what drives the break-up of continental plates.

Earth's mantle, which underlies the tectonic plates, moves slowly, rather like a liquid being warmed up. Earth's heat comes from the heat left over from processes that formed the planet and from the radioactive decay of elements such as uranium. Rising upward through this slowly circulating matter are columns of warm, buoyant material — mantle plumes — that effectively transport heat from great depths to the surface^{2,3}. These plumes are probably the source of large lava outpourings, such as those found in Siberia and India, that may have caused mass species extinctions in the past⁴. But unambiguous detection of mantle plumes has so far been difficult⁵. This is especially so for the detection of their deeper parts, which can be imaged only indirectly.

Furthermore, the effect of mantle plumes on Earth's surface is far from simple. Plumes are thought to have a symmetric shape that would generate a circular surface uplift once they impinge on the lithospheric plate. The Hawaiian plume, however, shows that the situation may be more complicated: horizontal plate motion, a possibly tilted plume conduit, small-scale convection in the upper mantle,

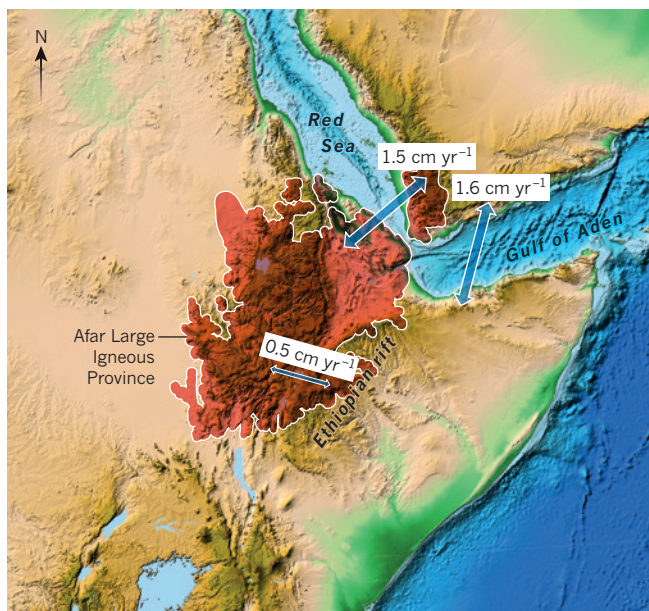


Figure 1 | Topography and bathymetry of the Afar region. Large amounts of volcanic rock erupted across Ethiopia and Yemen about 30 million years ago. These form the Afar Large Igneous Province and are usually thought to derive from the Afar mantle plume that underlies the region. Slightly before and after the volcanic event, rifting began in the Red Sea, Gulf of Aden and Ethiopia. Currently, plates in these rifts move away from each other by about 0.5 to 1.6 centimetres per year¹¹ (blue arrows). It is unclear whether this continental rifting caused the volcanism or whether the Afar plume helped the rifts to form. Burov and Gerya's computer model¹ could favour the latter possibility, in which continental rifting was localized above the Afar mantle plume. Land is shown in brown–green, and areas below sea level in blue. (Map made with GPlates software¹² using topographic and bathymetric data from the US National Geophysical Data Center¹³; outline of the Afar Large Igneous Province after ref. 14.)

and heterogeneities in the overlying plate all combine to destroy the expected surface symmetry^{6,7}.

The topography generated by mantle flow occurs over wide areas, but with low magnitudes — up to hundreds of metres of surface elevation over distances of hundreds to thousands of kilometres. Locally, this signal is easily overwhelmed by the larger surface displacements generated by tectonic processes such as mountain building. Topography caused by mantle flow is therefore usually determined by filtering Earth's topography from the effects of density variations caused by plate tectonics. But Burov and Gerya present three-dimensional (3D) computer models showing that

Earth's topography cannot always be separated into deep and shallow contributors when plate–mantle interactions come into play. This is particularly so for continental plates, which are thicker than oceanic plates and characterized by a horizontal layering of alternating strong and weak materials. Earlier 2D computer experiments⁸ showed that the continental lithosphere above a plume head can develop drip-like instabilities. Deformation along the weak layers inside the lithosphere may lead to alternating small-scale uplift and subsidence patterns, instead of a single, dome-shaped surface uplift. Using high-resolution models of plume–lithosphere interaction, Burov and Gerya demonstrate that this also holds for 3D models.

These models combining mantle flow with deformation of tectonic plates imply that the contributions from plate tectonics or mantle flow no longer need to be considered individually. Instead, combined 3D models predict a surface topography that can be directly compared with observations. Nonetheless, such comparisons may be difficult, because Burov and Gerya's results show that continental topography above

mantle plumes not only varies in space and time, but also can be asymmetric. It therefore becomes a challenge to identify surface uplift related to mantle plumes unambiguously, or even to use such surface observations to constrain properties of Earth's interior, such as viscosity.

Plume–lithosphere interaction models such as those of Burov and Gerya may contribute to the debate over a chicken-and-egg question in Earth science: can mantle plumes cause so much uplift and deformation in a continent that it breaks and creates a new ocean? Or do continents rift apart under stresses that are generated by tectonic-plate motions, and does the break-up process cause the upwelling

EVOLUTIONARY DEVELOPMENTAL BIOLOGY

of warm mantle from which the ocean crust is formed? As is often the case, the answer is probably a bit of both.

Continental break-up and the eruption of large amounts of magma are often closely related⁹. For example, the opening of the central Atlantic Ocean occurred shortly after the formation of the Central Atlantic Magmatic Province (about 200 million years ago), and magmatism and rifting in the Afar region of northeastern Africa are closely related in space and time (Fig. 1). In other places, however, continental break-up occurred without much magma, for example in the break-up between Iberia and Newfoundland. The fact that the processes occur at the same time does not favour one scenario over the other. It is here that Burov and Gerya's models offer new insight. The authors show that slow continental extension might not lead to break-up when continents do not contain heterogeneities inherited from earlier deformations, such as faults or changes in rock strength. But if a plume impinges on a continent that is already undergoing slow extension, it may localize deformation and help the plate to break up. The mantle plume alone does not cause continental break-up, but it could be the deciding factor.

Such a picture of plume-assisted rifting might apply to the separation of Norway and Greenland about 54 million years ago. Here, rifting events over a few hundred million years led to break-up only shortly after the North Atlantic Igneous Province formed. It would be interesting to see if a plume that is offset from a developing rift could lead to "upside-down drainage" flows¹⁰, in which plume material moves upward along the base of the plate towards the rift area. This would considerably increase the likelihood of plume–rift interactions. ■

Susanne Buiter is at the Geological Survey of Norway, 7040 Trondheim, Norway, and the Centre for Earth Evolution and Dynamics, University of Oslo, Norway.
e-mail: susanne.buiter@ngu.no

1. Burov, E. & Gerya, T. *Nature* **513**, 85–89 (2014).
2. Wilson, J. T. *Can. J. Phys.* **41**, 863–870 (1963).
3. Morgan, W. J. *Nature* **230**, 42–43 (1971).
4. Courtillot, V., Jaeger, J. J., Yang, Z., Féraud, G. & Hofmann, C. *Geol. Soc. Am. Spec. Pap.* **307**, 513–525 (1996).
5. Foulger, G. R. *Astron. Geophys.* **43**, 6.19–6.24 (2002).
6. Ballmer, M. D., Ito, G., van Hunen, J. & Tackley, P. J. *Nature Geosci.* **4**, 457–460 (2011).
7. Rychert, C. A., Laske, G., Harmon, N. & Shearer, P. M. *Nature Geosci.* **6**, 657–660 (2013).
8. Burov, E. & Guillou-Frottier, L. *Geophys. J. Int.* **161**, 469–490 (2005).
9. Courtillot, V., Jaupart, C., Manighetti, I., Tapponnier, P. & Besse, J. *Earth Planet. Sci. Lett.* **166**, 177–195 (1999).
10. Sleep, N. H. *J. Geophys. Res.* **102**, 10001–10012 (1997).
11. Nocquet, J.-M., Willis, P. & Garcia, S. J. *Geodesy* **80**, 591–607 (2006).
12. www.gplates.org
13. www.ngdc.noaa.gov
14. Coffin, M. F. & Eldholm, O. *Rev. Geophys.* **32**, 1–36 (1994).

Dynasty of the plastic fish

Ambitious experimental and morphological studies of a modern fish show how developmental flexibility may have helped early 'fishapods' to make the transition from finned aquatic animals to tetrapods that walk on land. [SEE ARTICLE P.54](#)

JOHN HUTCHINSON

Napoleon Bonaparte's military excursions into Egypt in 1798–99 led a young French naturalist, Étienne Geoffroy Saint-Hilaire, to cross paths with a strange fish that had paired lungs and could 'walk' on land on its stubby, lobe-like fins. In 1802, he dubbed this fish "*Polyptère bichir*", today known as the Nile bichir (*Polypterus bichir*). The bichir's melange of primitive and advanced traits helped to establish Geoffroy as a leading anatomist, embryologist and early evolutionary-development researcher of repute even today². Now, on page 54 of this issue, Standen *et al.*³ focus on *Polypterus* in their own excursion under the 'evo-devo' flag that Geoffroy helped to raise. The authors suggest that the remarkable plasticity of the skeleton of *Polypterus senegalus* (the smaller West African relative of *P. bichir*) reveals a key part of the mechanism that might have facilitated the gradual transition of limbed vertebrates from water to land.

In a bold experiment, Standen and colleagues reared a group of bichirs on land for eight months and compared them with bichirs that had developed in their normal aquatic environment (Fig. 1). They then studied how the fish from the two groups moved on land, and how the shape of the skeletal elements of their paired front fin bases differed. They found that, compared with water-raised bichirs, the land-acclimated fish took faster steps, their fins 'slipped' across the substrate less frequently, they held their fins closer to their bodies, their noses stayed more aloft and their tails undulated less, with less-variable motions overall. These were behaviours that the authors had predicted should develop to enhance walking abilities on land.

Furthermore, the bones of the neck and shoulder region in the land-reared fish had altered in shape to produce a more mobile fin base with greater independence of motion between the fin and the neck, along with improved bracing of the ventral 'collarbone' region. These environmentally induced traits probably fostered the locomotor changes observed in the land-reared fish and helped the animals to resist gravity, thereby representing

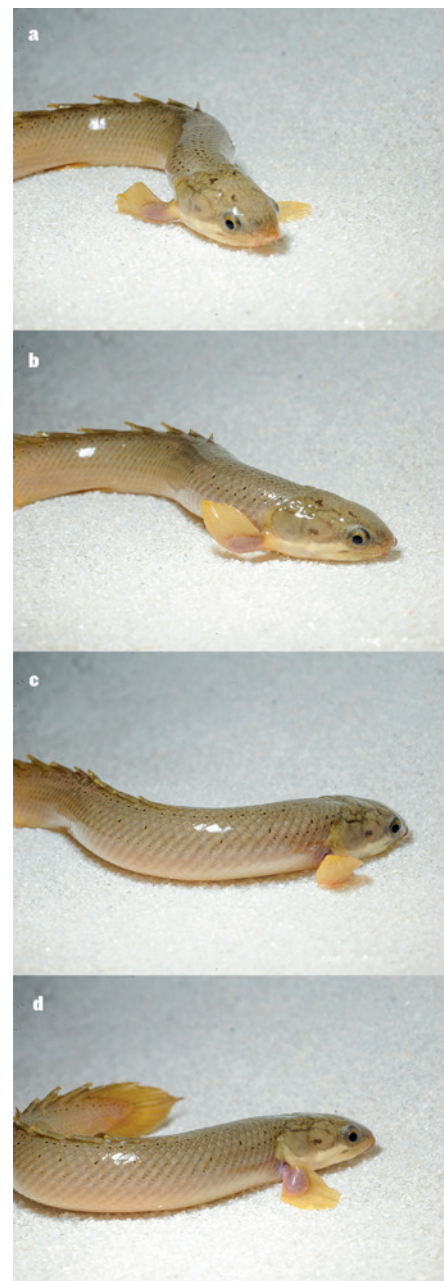


Figure 1 | Walk this way. The typical walking sequence of *Polypterus bichir* on land. **a**, The fish plants its left fin while its right fin swings forward; **b**, **c**, the head and tail turn towards the left fin; **d**, finally, the right fin is planted on the ground as the left is raised.

ANTOINE MORIN

a common biological phenomenon termed developmental plasticity^{4,5}. Surprisingly, the land-reared fish could swim just about as well as the aquatic cohort, so there was no clear trade-off between being a good swimmer and a good walker.

Considered alone, the developmental plasticity of bichir form and function shows how impressive these amphibious fish are. But Standen and colleagues ventured further, to apply the lessons learned from bichir ontogeny to a phylogenetic context and a macroevolutionary question.

The authors infer that the plasticity of bichir development could have been harnessed during the evolutionary transformation of fins used for swimming into limbs used for walking, in the 'fishapod' ancestors of tetrapods (four-limbed vertebrates). Indeed, bichirs are close to the base of the family tree of fishes⁶, and other living relatives of tetrapods have reduced or no fins (for example, lungfishes), or are adapted to strange deep-sea swimming lifestyles, never walking on land (coelacanths). Therefore, perhaps bichirs and the fishapod lineage share what Geoffroy called 'unity of type', today termed homology, with regard to their developmental plasticity in response to a land environment. Surveying the fossil record of early fishapods and tetrapods, Standen *et al.* found that the macroevolutionary changes of neck and shoulder anatomy in these gradually land-adapted animals parallel changes that they observed in 'terrestrialized' *Polypterus*, providing support for this hypothesis.

Further testing of the relevance of *Polypterus*'s plasticity to tetrapod evolution is, of course, difficult. However, the fishapod lineage has some exceptional examples of fossil preservation, and a rigorous indirect test of this hypothesis might be possible if there are sufficient sample sizes (for example, from fossil beds that reveal specimens at different developmental stages, such as the Late Devonian Miguasha site in Canada⁷) and palaeoenvironmental gradients in fish or tetrapod habitats. Even small samples could be helpful. For example, the early tetrapod *Ichthyostega* exhibited some developmental changes in its forelimb suggesting that individuals became more terrestrial as they grew, whereas the related *Acanthostega* did not show such changes⁸, hinting at developmental plasticity in the former animal.

Might it be that, during the Devonian period (around 360 million to 420 million years ago), the fishapod ancestors of tetrapods were intermittently floundering about on land, gradually shifting from anatomy and behaviours that were more developmentally plastic (as in bichirs) to ones that were more canalized into the forms and functions of land-adapted tetrapods? We don't know, but Standen and colleagues suggest that the developmental plasticity could have led to fixation (reduction of plasticity). This is an example of a

proposed evolutionary phenomenon called genetic assimilation — a concept promoted by Conrad Hal Waddington, an intellectual descendant of Geoffroy, from the 1950s onwards⁹. There is some empirical support for this idea¹⁰, and the findings in bichirs may eventually add to it.

The nature of the genetic and developmental mechanisms by which bichirs achieve developmental plasticity is unclear. If the plasticity is sufficiently heritable, then it might be selected for in multi-generational experiments, such that (with enough time and luck raising these unusual fish) we could directly test the hypothesis that the animals' plastic response to a terrestrial environment can become genetically assimilated. Such a study could thus become an exemplar of how genetic assimilation can contribute not only to microevolutionary change, but also to macroevolutionary events, as has been previously suggested⁴.

Geoffroy would probably have applauded Standen and colleagues' study of developmental plasticity, all the more for involving his beloved bichirs. Much as Napoleon's landfall in Egypt was not a lasting success, bichirs never produced wholly terrestrial descendants,

despite their malleable locomotor system. But the same type of plastic developmental mechanism that bichirs use today to make tentative, floppy incursions of the terrestrial realm might have been harnessed by our own fishapod forebears, leaving a much more revolutionary dynasty on Earth. ■

John Hutchinson is in the Structure and Motion Laboratory, Department of Comparative Biomedical Sciences, Royal Veterinary College, Hatfield AL9 7TA, UK. e-mail: jhutchinson@rvc.ac.uk

1. Geoffroy, E. *Ann. Mus. Hist. Nat.* **1**, 57–68 (1802).
2. Le Guyader, H. *Geoffroy Saint-Hilaire: A Visionary Naturalist* (transl. Grene, M.) (Univ. Chicago Press, 2004).
3. Standen, E. M., Du, T. Y. & Larsson, H. C. E. *Nature* **513**, 54–58 (2014).
4. West-Eberhard, M. J. *Annu. Rev. Ecol. Syst.* **20**, 249–278 (1989).
5. Pigliucci, M., Murren, C. J. & Schlichting, C. D. *J. Exp. Biol.* **209**, 2362–2367 (2006).
6. Near, T. J. *et al.* *Evolution* **68**, 1014–1026 (2014).
7. Cloutier, R. *Geosci. Can.* **40**, 149–163 (2013).
8. Callier, V., Clack, J. A. & Ahlberg, P. E. *Science* **324**, 364–367 (2009).
9. Waddington, C. H. *Evolution* **7**, 118–126 (1953).
10. Crispo, E. *Evolution* **61**, 2469–2479 (2007).

This article was published online on 27 August 2014.

CELL BIOLOGY

A guardian angel of cell integrity

The finding that RIPK1, an integral protein in cell–death pathways, also functions to preserve the body's epithelial–cell barriers challenges the idea that cell death and survival are regulated by distinct factors. SEE LETTERS P.90 & P.95

FRANCIS KA-MING CHAN

The saying, “while you do not know life, how can you know death?”, from the great Chinese philosopher Confucius, encapsulates the human desire to understand matters of life and death. Beyond philosophical curiosity, at the cellular level the balance between life and death is essential for health. Mutations in genes that regulate cell death have been found to cause human diseases ranging from inflammation and autoimmunity to cancer. Normally, cell-death genes have dedicated roles in either cell survival or cell death. But in this issue, two groups^{1,2} show that receptor interacting protein kinase 1 (RIPK1), a factor known to promote cell death, has a paradoxical function in supporting the survival of the epithelial cells that line our body's cavities and skin surfaces.

Since the discovery of RIPK1 almost 20 years ago³, the proposed function of the protein

has alternated between that of a cell-death promoter and a cell-death inhibitor. Much of what we know about RIPK1 biology comes from studies of its activation by the tumour-necrosis factor (TNF) receptor, a cell-surface protein that, when bound by TNF molecules, contributes to inflammation, cell differentiation and cell death. Genetically modified mice that lack RIPK1 die shortly after birth^{4–6}, and this was thought to be due to failure to stimulate NF- κ B, a cellular factor that turns on the expression of survival genes. Consistent with this idea, extensive cell death and inflammation in multiple tissues occur in these RIPK1-deficient mice.

But, in contrast to this survival function, RIPK1 also binds to and activates FADD and caspase 8 — two members of a protein complex that triggers a form of cell death called apoptosis⁷. To complicate matters further, RIPK1 can also partner with its cousin RIPK3 to induce an inflammatory form of cell

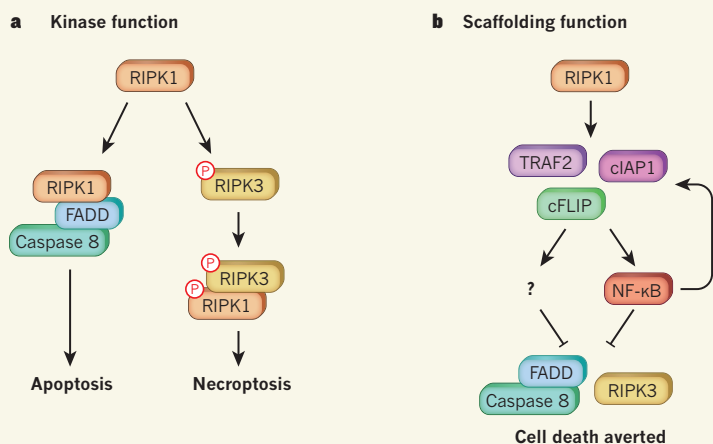


Figure 1 | Regulation of survival and death by RIPK1. **a**, The protein RIPK1 drives cell death through apoptosis, by activating the FADD–caspase-8 protein complex, and through necroptosis, which involves interaction between phosphorylated forms of RIPK1 and its partner protein RIPK3. These death-inducing activities of RIPK1 require the protein's activity as a kinase enzyme. **b**, But RIPK1 can also promote cell survival. Dannappel *et al.*¹ and Takahashi *et al.*² show that, in epithelial cells in the gut, this occurs independently of the protein's kinase activity and involves maintaining the activity and levels of the proteins TRAF2, cIAP1 and cFLIP — possibly through a 'scaffolding' process in which RIPK1 shields these proteins from degradation. This function may involve the NF-κB signalling pathway, but also seems able to act by an NF-κB-independent mechanism.

death termed necroptosis⁸. These apparently conflicting functions of RIPK1 have puzzled researchers for years, and we are still struggling to join the dots on how RIPK1 deficiency leads to perinatal lethality.

To try to fill this gap, Dannappel *et al.*¹ and Takahashi *et al.*² generated mice lacking RIPK1 specifically in the intestinal epithelium. These mice survived longer after birth than mice with whole-body deficiency of RIPK1, but they developed severe inflammatory bowel disease within the first few weeks of life and died as a result. This colitis was largely due to the heightened sensitivity of the mice to TNF. Histological examination revealed that the intestinal epithelial cells in these animals had undergone extensive apoptosis.

In contrast to previous reports^{9,10}, the activity of the survival factor NF-κB was not significantly affected in the RIPK1-deficient cells. Hence, the sensitivity of these cells to death was not a result of impaired NF-κB. Previous work had shown^{11,12} that the function of RIPK1 as a kinase enzyme is essential for cell death by apoptosis and necroptosis, but not for its stimulation of NF-κB. However, both Dannappel *et al.* and Takahashi *et al.* found that the kinase function of RIPK1 is also not responsible for the severe colitis, because mice engineered to express a kinase-inactive version of RIPK1 developed normally and showed no abnormalities. Nonetheless, the severe colitis was fully reversed when FADD–caspase-8-mediated apoptosis and RIPK3-dependent necroptosis were both inactivated. These results indicate that a yet-to-be defined, kinase-independent function of RIPK1 is responsible for protecting intestinal cells from injury.

Dannappel and colleagues also found that deletion of RIPK1 only in the skin epidermis led to psoriasis-like inflammation. However, unlike in the intestine, blocking necroptosis alone was sufficient to prevent this skin inflammation. Inhibition of FADD–caspase-8-mediated apoptosis and RIPK3-dependent necroptosis has also previously been found to prevent the perinatal lethality of mice with whole-body RIPK1 deficiency^{4–6}. Together, these results indicate that RIPK1 promotes cell survival by inhibiting apoptosis and necroptosis and, hence, that the protein has the enigmatic role of both a promoter and an inhibitor of cell death (Fig. 1).

How does RIPK1 orchestrate these diametrically opposing signals? Both groups found that, on TNF stimulation, RIPK1-deficient cells lost expression of the proteins cIAP1, TRAF2 and cFLIP — proteins that switch on the pro-survival factor NF-κB. Moreover, active NF-κB can further increase the levels of these factors. Thus, RIPK1 may protect cells by preserving the integrity of survival proteins such as cIAP1, TRAF2 and cFLIP. In this regard, it is noteworthy that these proteins are regulated by ubiquitination, a process that often tags proteins for degradation by a macromolecular structure called the proteasome. Because the kinase function of RIPK1 is not required for cell survival, it is possible that RIPK1 forms a protective 'scaffold' that shields the survival proteins from ubiquitination and proteasomal degradation. Alternatively, RIPK1 may directly inhibit FADD–caspase-8-mediated apoptosis and RIPK3-dependent necroptosis.

Regardless of the molecular mechanism



50 Years Ago

'Interferometric Raman spectroscopy using infra-red excitation' — Many pure substances and industrial intermediates are strongly coloured and their Raman spectra cannot be recorded using ultra-violet and visible lines, but nearly all compounds have a region of transparency in the near infra-red ... Our sample tube used with the laser [has] plain glass windows at each end so that the exciting radiation passes through the sample and directly into the interferometer. The laser is pulsed and the photomultiplier voltage pulses are 'gated' to remove the noise between pulses. The output is recorded digitally ... and spectra are obtained by Fourier transformation in a digital computer ... Using this apparatus we have observed the Raman spectrum of iodine dissolved in carbon tetrachloride at a molar ratio of 1:1,000 ... The possibilities of Raman measurements with near infra-red exciting lines have thus been demonstrated.

From *Nature* 5 September 1964

100 Years Ago

'The type-reading optophone' — Any instrument designed for translating optical into acoustic effects, or light into sound, and thus to some extent substituting the ear for the eye, may be appropriately termed an "optophone" ... The latest of these ... is designed with the object of enabling blind persons to "read" ordinary letterpress by means of the ear ... An optical system throws the image of a glowing Nernst filament upon the printed paper, laid face downwards on a suitably perforated desk. This image is broken up into a series of ... luminous dots, flashing with different musical frequencies, by means of a rotating siren disc ... It has been found possible to obtain a "readable" sound from type of the ordinary newspaper size.

From *Nature* 3 September 1914

involved, it is peculiar that, in RIPK1-deficient mice, embryonic development is normal and cell injury and inflammation manifest themselves only at birth. What might be the physiological cue that triggers these events after birth? Newborns encounter many changes in their environment, including altered oxygen levels, exposure to ambient light and colonization of the intestine by commensal bacteria. This bacterial colonization can strongly influence the development of immunity¹³. Surprisingly, although Takahashi *et al.* showed that antibiotic treatment to eliminate commensal bacteria protected RIPK1-deficient mice from lethal colitis, Dannappel *et al.* found that colitis still developed in antibiotic-treated

mice and in germ-free mice. The reason behind these discrepant observations is not known. Moreover, although TNF contributes significantly to the colitis in RIPK1-deficient mice, it is not the only driver of the disease. It now remains for biologists to take up the gauntlet to decipher these puzzles and to identify the physiological cues and mechanisms that trigger RIPK1-dependent survival responses. ■

Francis Ka-Ming Chan is in the Department of Pathology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA.
e-mail: francis.chan@umassmed.edu

1. Dannappel, M. *et al.* *Nature* **513**, 90–94 (2014).
2. Takahashi, N. *et al.* *Nature* **513**, 95–99 (2014).
3. Stanger, B. Z., Leder, P., Lee, T.-H., Kim, E. & Seed, B. *Cell* **81**, 513–523 (1995).
4. Rickard, J. A. *et al.* *Cell* **157**, 1175–1188 (2014).
5. Dillon, C. P. *et al.* *Cell* **157**, 1189–1202 (2014).
6. Kaiser, W. J. *et al.* *Proc. Natl Acad. Sci. USA* **111**, 7753–7758 (2014).
7. Edinger, A. L. & Thompson, C. B. *Curr. Opin. Cell Biol.* **16**, 663–669 (2004).
8. Moriwaki, K. & Chan, F. K.-M. *Genes Dev.* **27**, 1640–1649 (2013).
9. Kelliher, M. A. *et al.* *Immunity* **9**, 297–303 (1998).
10. Ting, A. T., Pimentel-Muñoz, F. X. & Seed, B. *EMBO J.* **15**, 6189–6196 (1996).
11. Chan, F. K.-M. *et al.* *J. Biol. Chem.* **278**, 51613–51621 (2003).
12. Wang, L., Du, F. & Wang, X. *Cell* **133**, 693–703 (2008).
13. Hooper, L. V., Littman, D. R. & Macpherson, A. J. *Science* **336**, 1268–1273 (2012).

with today's technology, but understanding what any changes mean can be even harder than reading certain passages by Joyce: some genetic text might as well be written in Elvish. At the next level up, and with some noteworthy exceptions, we have little understanding of which parts of the proteins and RNA molecules encoded by genes matter to the cell and the organism, and which do not.

This seems a simple enough problem to solve. To discover whether variations in a gene matter, alter them and see what happens. However, this experiment has long proved challenging, because our cells have evolved to protect their DNA from such change on demand. The solution is gene editing², the first step of which is to cut DNA inside the living cell. This was initially achieved using an engineered enzyme called a zinc-finger nuclease³, and more recently, two other types of enzyme — TALENs⁴ and CRISPR-associated enzymes⁵ — have also been co-opted for this purpose. Cells can repair DNA breaks by either joining the two ends back together or patching up the break using genetic information from a different DNA molecule that has an identical or similar sequence. In one form of human-gene editing⁶, this second pathway is subverted, by having the cell repair the break using a new piece of DNA that contains a desired mutation.

Findlay *et al.* built on this latter approach to address a fundamental question: which parts of a gene are actually useful? First, they focused on a gene whose function is vital for cell survival, *DBR1*, engineering a stretch of 75 nucleotides such that every possible single-nucleotide mutation was made individually. Genome editing has previously been used to create a small set of desired changes in a mammalian gene^{6,7}, or to edit multiple genes in the same pathway^{8,9}, but never before has every change possible been made (until now, experiments of such scale were achievable only in budding yeast¹⁰). The result is a group of cells that each carry a different DNA sequence, much like a set of sentences that each differ by one typographical error (Fig. 1).

BIOLOGICAL TECHNIQUES

Edit the genome to understand it

Genome editing followed by sequencing has now been used to engineer and analyse every variation of several stretches of human DNA in living cells, providing insight into the function of each constituent nucleotide molecule. [SEE LETTER P.120](#)

FYODOR D. URNOV

You were shaped by both nature and nurture, by your DNA and your environment. How do the differences in your genome, compared with those of other people, affect who you are? In this issue, Findlay *et al.*¹ (page 120) put a combination of human-genome editing² and in-depth gene sequencing to brilliant use, and bring us one large step closer to a general answer to this monumental question.

Understanding our genes is a challenge. Human DNA is around 6.5 billion base pairs long — as many letters as there are in more than 5,000 copies of James Joyce's *Ulysses* — and differences in the DNA of any two people occur approximately every 1,000 base pairs. Although some of these differences have a large impact (for example, preventing certain adults from drinking milk), most have either a minor effect (such as making one person 0.5 centimetres taller than another) or none at all. DNA can be read quickly and inexpensively

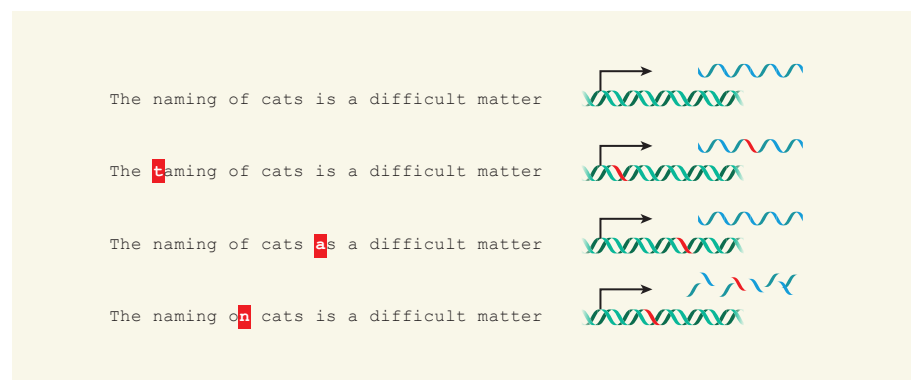


Figure 1 | The search for meaning. Findlay *et al.*¹ changed every nucleotide molecule in a short stretch of DNA, and then analysed the resulting mutant cells to determine which changes were beneficial and which harmful. This can be compared to generating and then reading a set of sentences in which one letter is changed each time (such as these, based on a poem by T. S. Eliot). Some changes alter the meaning (in DNA, this might affect what the protein encoded by the gene does), some are innocuous, and some make the sentence nonsensical (in which case, the RNA transcribed from the gene may be destroyed). This informs us about the meaning and role of words that comprise the sentence, or of nucleotides in a DNA sequence.

To determine which changes are beneficial to the cell and which detrimental, Findlay and colleagues used deep sequencing, which reads every copy of every gene in every cell of a population. Immediately after editing, the cells are a kaleidoscope of genetic diversity. Edited cells account for only 1–3% of the total cell population (lower than seen in other studies²), but this is not a real problem because deep sequencing can identify even very rare DNA sequences.

After a few days, a stark change occurs. Many new sequences disappear or diminish in number. This is survival of the fittest at the cellular level. The authors found that cells unlucky enough to acquire a change in a nucleotide needed for gene function died immediately, but cells that had more-benign errors lived on. This experiment provides a remarkable functional map of this bit of genetic text — we know whether each and every position makes a useful contribution to the working of the whole protein.

Some genetic changes do not affect what a protein does, but rather how messenger RNA molecules are put together such that sections that do not specify the sequence of a protein are removed (a process known as splicing). Findlay and co-workers investigated how DNA sequence affects splicing in *BRCA1*, mutations in which cause breast cancer, in some cases because of improper splicing.

The authors generated almost every possible sequence in a 6-base-pair stretch of *BRCA1*, and investigated which sequences helped the gene to be copied into normal RNA, and which prevented it. They took this remarkable group of 4,048 different kinds of cell, growing side by side in the same Petri dish, and measured how often each sequence occurred in *BRCA1* DNA and the corresponding RNA. Some sequences were never found in RNA, giving an insight into which genetic signals control how RNA acquires its fully functional form.

Findlay and colleagues have provided a way to find meaning in the text of human DNA, by systematically analysing each nucleotide in a gene in its normal setting in the chromosome. All you need is a robust way to edit your region of interest^{3–5} and a method to assay the cellular consequences of editing. The word ‘random’ often has negative connotations in science, but not in this instance. Making random changes in a gene and letting nature take its course is enormously informative. For instance, a major challenge for women who carry a mutation in *BRCA1* is to determine the risk of contracting cancer for their specific mutation. Findlay and co-workers’ approach can be used to address this problem and to determine which specific *BRCA1* mutations are the most worrisome.

More generally, the juxtaposition of genome editing and deep-sequencing technologies will, without doubt, provide a basis for progress in our quest to understand how our DNA makes

us who we are. The authors’ work provides an excellent case in point to support the words of geneticist Sydney Brenner¹¹: “Progress in science results from new technologies, new discoveries and new ideas, probably in that order.” ■

Fyodor D. Urnov is at Sangamo BioSciences Inc., Richmond, California 94804, USA.
e-mail: furnov@sangamo.com

1. Findlay, G. M., Boyle, E. A., Hause, R. J., Klein, J. C. & Shendure, J. *Nature* **513**, 120–123 (2014).
2. Carroll, D. *Annu. Rev. Biochem.* **83**, 409–439 (2014).

3. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. *Nature Rev. Genet.* **11**, 636–646 (2010).
4. Joung, J. K. & Sander, J. D. *Nature Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
5. Ran, F. A. et al. *Nature Protocols* **8**, 2281–2308 (2013).
6. Urnov, F. D. et al. *Nature* **435**, 646–651 (2005).
7. Goldberg, A. D. et al. *Cell* **140**, 678–691 (2010).
8. Doyon, J. B. et al. *Nature Cell Biol.* **13**, 331–337 (2011).
9. Sexton, A. et al. *Genes Dev.* <http://dx.doi.org/10.1101/gad.246819.114> (2014).
10. Braberg, H. et al. *Cell* **154**, 775–788 (2013).
11. Robertson, M. *Nature* **285**, 358–359 (1980).

This article was published online on 20 August 2014.

COSMOLOGY

Meet the Laniakea supercluster

An analysis of a three-dimensional map of galaxies and their velocities reveals the hitherto unknown edges of the large system of galaxies in which we live — dubbed the Laniakea supercluster. [SEE LETTER P.71](#)

ELMO TEMPEL

One of the greatest advances in cosmology has been the discovery of how matter and light are organized on scales larger than those of galaxies^{1,2}. However, despite tremendous effort, astronomers have been unable to map in detail the large-scale cosmic structure in which the Milky Way resides. Now, on page 71 of this issue, Tully et al.³ report an analysis of data from a vast

catalogue of galaxies that has allowed them to do just that.

The large-scale structure of the Universe is an intricate network of clusters, filaments and superclusters of galaxies, together with cosmic voids that are almost empty of galaxies. Superclusters are extended regions containing a large number of galaxies, but this concept is rather vague; researchers lack a robust, quantitative definition for them. Tully and colleagues have found a neat way of identifying the edges of

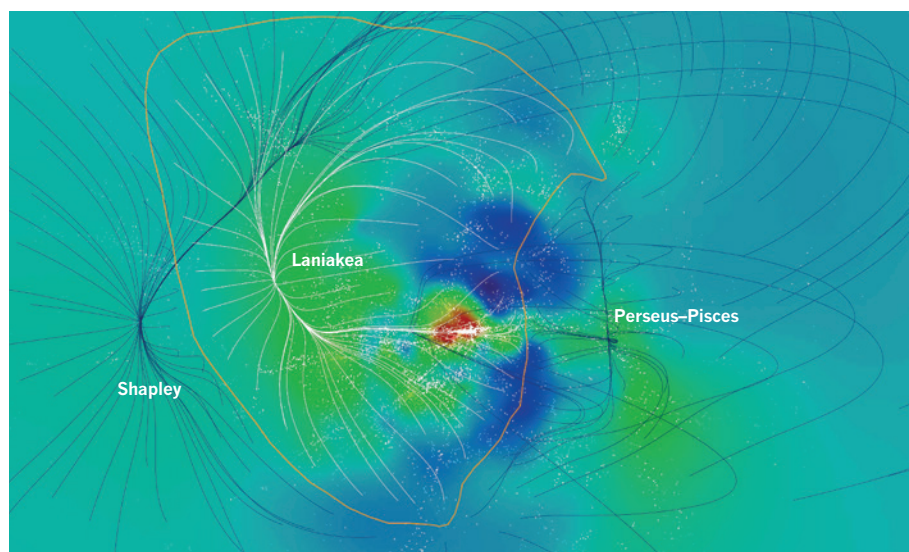


Figure 1 | The edges of our home supercluster of galaxies. The image shows a slice through the Laniakea supercluster and adjacent Shapley and Perseus–Pisces superclusters, as identified by Tully and colleagues³. Areas of high galaxy density are shown in red and cosmic voids in dark blue. The Milky Way galaxy lies essentially in a plane parallel to the slice in the centre of the figure. Velocity streams along which galaxies move within our supercluster are shown in white, with other velocity streams in dark blue.

REF. 3

superclusters by examining the motions of galaxies. In doing so, they have detected the boundaries of our home supercluster, which they have called the Laniakea supercluster. Their paper is supplemented by a beautiful movie (<http://irfu.cea.fr/laniakea>) that shows our supercluster and its dynamical connection to other neighbouring large-scale systems. The movie is essential for comprehending the complexity of cosmic structures.

Mapping the large-scale structure of the nearby region of the Universe is important for several reasons. First, it reveals details of the large-scale cosmic structures that surround the Milky Way. These details are nearly impossible to observe for systems far away from Earth. Second, the morphology of the nearby Universe is essential for a precise determination of cosmological parameters such as the density of dark energy⁴, which is thought to drive the acceleration of the expanding Universe. Third, examination of cosmic structures around the Milky Way will help us to understand how the Galaxy formed and evolved⁵, and galaxy formation processes in general.

Tully and colleagues' study is based on data from the Cosmicflows-2 galaxy catalogue⁶. The authors combined existing measurements of the velocities at which galaxies recede from Earth — which are mainly caused by the cosmic expansion and provide an indirect estimate of how far away they are — with direct galaxy distance measurements from the Cosmicflows-2 data set. This enabled them to derive the 'peculiar velocities' of the galaxies, that is, their true velocity relative to a rest frame. The peculiar velocity is obtained by subtracting the contribution of the cosmic expansion, which is determined using the direct distance measurement, from the recession velocity.

Direct distance measurements of galaxies are extremely difficult to perform, and the lack of such data has limited this kind of analysis in the past. However, the use of peculiar velocities can provide information about cosmic structures that is otherwise hard to obtain. And in the present case, it allowed the extent, structure and dynamics of Earth's supercluster, as well as those of other nearby superclusters, to be determined. We can only imagine what other details and structures might be uncovered if additional direct-distance measurements of galaxies are carried out.

A noteworthy aspect of Tully and colleagues' study is the use of Wiener filtering⁷ — a nifty algorithm that translates an incomplete map of peculiar velocities of galaxies into a complete map of the underlying distribution (density field) and dynamics (velocity flow field) of matter (Fig. 1). It is this technique that allowed the authors to come up with a quantitative definition of a supercluster. According to their definition, a supercluster is a 'basin of attraction' in the velocity flow field. In other words, the boundaries of a supercluster are defined

by the places at which the velocity flow field points in different directions on either side of the boundary. This is the first clear definition of a supercluster. The downside of it is that it requires dynamical information that is available only for the nearby Universe.

Tully *et al.* find several basins of attraction in our corner of the Universe, including Laniakea and the previously known Perseus–Pisces and Shapley superclusters. Laniakea has a diameter of 160 million parsecs (520 million light years), and is much bigger than already identified superclusters in our local neighbourhood. However, it is smaller than the largest superclusters that have been found in the more distant Universe⁸. It is a surprise that Laniakea was not spotted in previous astronomical surveys. It seems that measurements of the peculiar velocities of galaxies are essential for identifying the boundaries of some superclusters.

Of course, these results do not mark the end of mapping the Universe. Although Tully *et al.* used the best galaxy catalogue available, these data do not extend far enough in cosmic space to explain the motion of our Galaxy with respect to the rest frame of the cosmic microwave background — relic radiation from the Big Bang. The Universe must be mapped on a much bigger scale than that achieved here to

fully understand what processes affected the formation of cosmic structures in our local Universe. This is a challenging task, but one that is worthwhile and that we must hope will be tackled using future surveys.

Finally, I praise the choice of the name Laniakea for Earth's supercluster. It is taken from the Hawaiian words *lani*, which means heaven, and *akea*, which means spacious or immeasurable. That is just the name one would expect for the whopping system that we live in. ■

Elmo Tempel is in the Department of Cosmology, Tartu Observatory, Tõravere 61602, Estonia.

e-mail: elmo.tempel@to.ee

1. Jöeveer, M., Einasto, J. & Tago, E. *Mon. Not. R. Astron. Soc.* **185**, 357–370 (1978).
2. Bond, J. R., Kofman, L. & Pogosyan, D. *Nature* **380**, 603–606 (1996).
3. Tully, R. B., Courtois, H., Hoffman, Y. & Pomarède, D. *Nature* **513**, 71–73 (2014).
4. Sinclair, B., Davis, T. M. & Haugbølle, T. *Astrophys. J.* **718**, 1445–1455 (2010).
5. Shaya, E. J. & Tully, R. B. *Mon. Not. R. Astron. Soc.* **436**, 2096–2119 (2013).
6. Tully, R. B. *et al. Astron. J.* **146**, 86 (2013).
7. Zaroubi, S., Hoffman, Y. & Dekel, A. *Astrophys. J.* **520**, 413–425 (1999).
8. Liivamägi, L. J., Tempel, E. & Saar, E. *Astron. Astrophys.* **539**, A80 (2012).

STRUCTURAL BIOLOGY

How fluorescent RNA gets its glow

Fluorescent tags are proving invaluable for tracking RNA molecules in cells. Two sets of crystal structures for one such tag — an RNA motif that fluoresces when bound to a dye — will aid the development of even better markers.

WILLIAM G. SCOTT

Green fluorescent protein is widely used as a visualization marker for biological molecules, and has revolutionized microscopic imaging in biological systems — a fact celebrated by the award of the 2008 Nobel Prize in Chemistry¹. Engineered fluorescent RNAs are potentially equally useful, and a green fluorescent RNA motif² called Spinach has been developed for this purpose. Uncovering the structural basis for how fluorescent RNAs work is crucial to realizing their full potential as experimental tools. Two sets of crystal structures of Spinach — one reported by Huang *et al.*³ in *Nature Chemical Biology*, and the other by Warner *et al.*⁴ in *Nature Structural and Molecular Biology* — now provide a deeper understanding of how it fluoresces, and should enable the design of improved labels for visualizing

individual RNA molecules in cells.

Fluorescence occurs when light shone on a molecule is absorbed, exciting the molecule, and is then re-emitted. The energy of the emitted light is lower than that absorbed, so a molecule excited by invisible ultraviolet light, for example, may fluoresce as highly visible green light. Because the fluorescent light is emitted in every direction, it can be measured at 90° from the direction of the light used to excite the molecule. Taken together, these effects can produce a highly sensitive signal with little background noise, potentially allowing the detection of just one or a very few molecules in a cell.

Proteins are not normally fluorescent. Green fluorescent protein (GFP), however, is an unusual enzyme that catalyses the chemical rearrangement of some of its own amino-acid side chains, creating an embedded molecule — known as a fluorophore — that absorbs

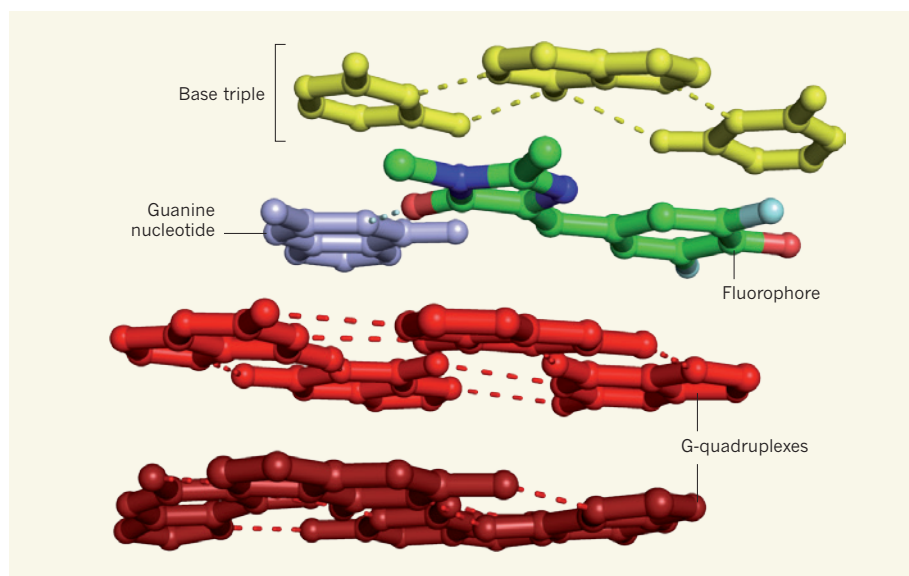


Figure 1 | Spinach sandwich. Two studies^{3,4} report crystal structures depicting how the ‘Spinach’ RNA motif binds its fluorophore — the dye molecule that fluoresces only when bound to Spinach. The structures reveal that the fluorophore binds tightly between a base triple (a structure formed from three nucleotide bases) and two stacked G-quadruplexes (each quadruplex is a coplanar duo of unusual RNA G–G base pairs; G is the nucleic-acid base guanine). The positioning and orientation of the fluorophore are further mediated by a coplanar guanine nucleotide. Broken lines indicate hydrogen bonding. In the fluorophore, oxygen atoms are shown in red, nitrogens in blue, fluorines in cyan and carbons in green. The figure was produced from coordinates for Huang and colleagues’ structure³, using PyMOL software (version 1.7.0.3).

ultraviolet light and fluoresces as bright green light.

Numerous complex cellular processes are controlled and orchestrated by RNA molecules, rather than by proteins. A particularly noteworthy example is RNA interference, in which small RNAs regulate, interfere with or inhibit gene expression. Furthermore, genes are expressed through the intermediate action of messenger RNA, which may be compartmentalized in a cell. The ability to tag and track the intracellular movement of any RNA by means of fluorescent molecules would therefore be of obvious use to biologists.

Unfortunately, nature has not provided a potential RNA tool analogous to GFP. Instead, using a process called *in vitro* directed evolution, biologists can identify RNA motifs that bind to small fluorescent molecules; these molecules are chemically similar to the fluorescent component of GFP and have similar fluorescent properties. Spinach is the most useful of such motifs, and can be fused to many RNAs of interest.

Spinach binds with high affinity to a synthetic dye molecule that resembles GFP’s fluorophore. The dye has the invaluable property of becoming fluorescent only when it binds to Spinach, and the further merit (as does GFP) of being non-toxic to cells. The fluorophore thus becomes visible only when it is bound to the RNA and illuminated with ultraviolet light, making it an ideal visualization marker.

The two sets of crystal structures for Spinach reveal a previously unknown fold and

fluorophore-binding site — the complexity of which defied prediction by computer programs commonly used to calculate RNA secondary structures. Huang *et al.* obtained their set of structures using an in-house approach⁵ in which the RNA was co-crystallized with an antibody. To address the potential criticism that the highly unusual RNA structure might be an artefact resulting from this method, the authors devoted considerable time and effort to providing many reassuring experimental controls.

The fluorophore can exist as four potential isomers, each of which can have multiple binding modes to Spinach. To identify the orientation of the bound fluorophore unambiguously, Huang and colleagues solved the crystal structure of the fluorophore alone, and that of the RNA bound to a bromine-bearing analogue of the fluorophore. The X-ray-absorption properties of the bromine allowed the binding position of the analogue, and therefore that of the original fluorophore, to be pinpointed. Huang and colleagues’ heroic undertaking has been unambiguously validated by the subsequent publication of Warner and co-workers’ crystal structures, which were obtained using a different (and more standard) crystallization approach.

So what have we learnt from the two sets of structures? Most importantly, the key to understanding how green fluorescent RNA works has been revealed. The fluorophore sits on a platform of two stacked G-quadruplexes (each quadruplex is a coplanar duo of unusual RNA G–G base pairs; G is guanine, a

nucleic-acid base). G-quadruplexes are often found at the ends of DNA molecules, but are seldom observed in RNA structures. The fluorophore is sandwiched tightly between the quadruplex platform and a coplanar RNA base triple (a structure analogous to a base pair, but involving three bases; Fig. 1). The binding pocket thus created enforces planarity on the bound fluorophore.

Two interrelated structural effects seem to be responsible for activating fluorescence. First, a negative charge on the oxygen atom attached to the fluorophore’s benzene ring is required for fluorescence. That negative charge is stabilized by RNA interactions in the binding site. These include hydrogen bonding to a nearby ribose structure; a ‘stacking’ interaction formed with the base triple that caps the binding site; and an electrostatic interaction with a nearby bound potassium ion that has a positive charge balancing the negative charge. Second, the large planar surface formed by the G-quadruplex platform provides an opportunity for extensive stacking interactions that greatly enhance fluorescence.

Our understanding of macromolecular structure and function can be put to the test by attempting to design molecules with a given function. Warner and colleagues demonstrated this by developing an improved green fluorescent RNA motif using the insight gleaned from their crystal structures. The resulting molecule is smaller and folds more efficiently than Spinach, and has been dubbed “Baby Spinach” by the authors. It is an ideal candidate marker for the next generation of RNA visualization experiments. ■

William G. Scott is in the Department of Chemistry & Biochemistry and the Center for the Molecular Biology of RNA, University of California, Santa Cruz, Santa Cruz, California 95064, USA. e-mail: wgsco@ucsc.edu

1. www.nobelprize.org/nobel_prizes/chemistry/laureates/2008/press.html
2. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. *Science* **333**, 642–646 (2011).
3. Huang, H. *et al. Nature Chem. Biol.* <http://dx.doi.org/10.1038/nchembio.1561> (2014).
4. Warner, K. D. *et al. Nature Struct. Mol. Biol.* <http://dx.doi.org/10.1038/nsmb.2865> (2014).
5. Koldobskaya, Y. *et al. Nature Struct. Mol. Biol.* **18**, 100–106 (2011).

CORRECTION

The News & Views article ‘Palaeoanthropology: The time of the last Neanderthals’ by William Davies (*Nature* **512**, 260–261; 2014) incorrectly named the modelled overlap period between Neanderthals and modern humans as 470–4,900 years (25–250 generations) instead of 2,600–5,400 years (130–270 generations).

Migrations and dynamics of the intertropical convergence zone

Tapio Schneider^{1,2}, Tobias Bischoff^{1,2} & Gerald H. Haug²

Rainfall on Earth is most intense in the intertropical convergence zone (ITCZ), a narrow belt of clouds centred on average around six degrees north of the Equator. The mean position of the ITCZ north of the Equator arises primarily because the Atlantic Ocean transports energy northward across the Equator, rendering the Northern Hemisphere warmer than the Southern Hemisphere. On seasonal and longer timescales, the ITCZ migrates, typically towards a warming hemisphere but with exceptions, such as during El Niño events. An emerging framework links the ITCZ to the atmospheric energy balance and may account for ITCZ variations on timescales from years to geological epochs.

The ITCZ can be identified as a tropical belt of deep convective clouds¹, or as the maximum in time-mean precipitation² (Fig. 1a). Seasonally, it migrates towards a hemisphere that warms relative to the other. Over the central Atlantic and Pacific oceans, the ITCZ migrates between 9° N in boreal summer and 2° N in boreal winter (Fig. 2a). Over the Indian Ocean and adjacent land surfaces, the ITCZ swings more dramatically between average latitudes of 20° N in boreal summer and 8° S in boreal winter, prompting the seasonal rainfall variations of the South Asian monsoon (Fig. 2b)³. On longer timescales, palaeo-records indicate that the ITCZ migrates similarly towards a differentially warming hemisphere^{4,5}. For example, the boreal-summer ITCZ appears to have migrated southward

as Northern Hemisphere summers cooled from the Holocene thermal maximum around 8,000 years before present (8 kyr BP) until the beginning of the Little Ice Age (LIA) at about 500 BP^{6–8}. Simultaneously, summer monsoon rainfall in parts of South Asia weakened (Fig. 3a)^{9–11}, either because the ITCZ swung less far northward in boreal summer, or because its rainfall intensity weakened, or a combination of both^{12,13}. These ITCZ variations over the Holocene arose because summer insolation in the Northern Hemisphere weakened with the precession of Earth's perihelion from Northern Hemisphere towards Southern-Hemisphere summer^{14,15}. But the annual-mean position of the ITCZ in the Northern Hemisphere (Fig. 1a) and the non-sinusoidal seasonal migrations in the South Asian monsoon sector (Fig. 2b)

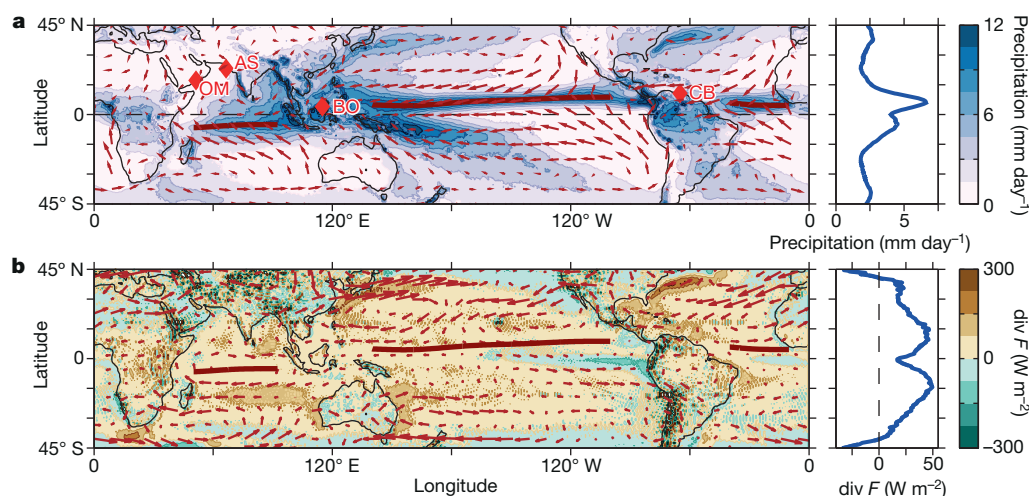


Figure 1 | Annual-mean precipitation, surface winds, and atmospheric energy balance. **a**, Precipitation (colour scale) is maximal at the ITCZ, and surface winds (vectors) converge there. The ITCZ over oceans (precipitation maxima) is marked by red lines. (Within the resolution of the data, convergence maxima of the surface winds and precipitation maxima coincide.) Other convergence zones, such as the South Pacific convergence zone, are also recognizable as precipitation maxima. The right panel shows the zonal-mean precipitation. The precipitation data are from the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA)⁶² for 1998–2012. The wind data are from the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis⁴⁴ for the same years. The longest wind vector (over the central Pacific) corresponds to a wind speed of

8.1 m s⁻¹. Also marked (red diamonds) are the sites of the palaeo-records shown in Figs 3 and 6: Oman (OM; 17° N, 51° E), Arabian Sea (AS; 23° N, 67° E), Borneo (BO; 4° N, 115° E), and the Cariaco basin (CB; 11° N, 65° W). **b**, Atmospheric moist static energy fluxes F (vectors) are weak near the ITCZ, but their divergence $\text{div } F$ (colour scale) is generally positive there. (The ITCZ over oceans is marked by the same red lines as in **a**.) The right panel shows the zonal-mean divergence of the moist static energy flux. The energy flux data are from the ECMWF interim reanalysis⁴⁴ for 1998–2012, corrected as in ref. 37 and provided by the National Center for Atmospheric Research. The longest vector (over the Kuroshio region) corresponds to a $2.7 \times 10^{10} \text{ W m}^{-1}$ energy flux.

¹California Institute of Technology, Pasadena, California 91125, USA. ²Department of Earth Sciences, ETH Zürich, 8092 Zürich, Switzerland.

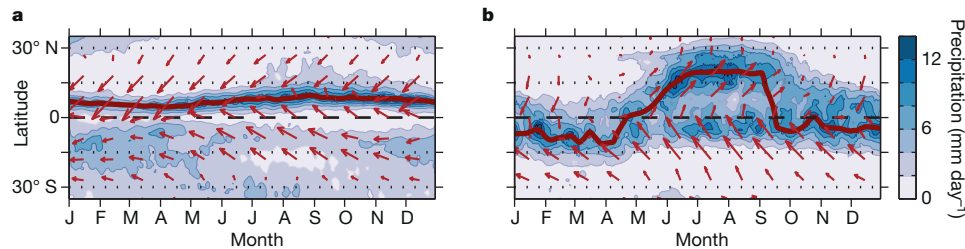


Figure 2 | Seasonal migration of the ITCZ over the Pacific and in the South Asian monsoon sector. Mean precipitation (colour scale) and surface winds (vectors) as a function of time of year averaged zonally over the Pacific (160° E–100° W) (a) and the South Asian monsoon sector (65° E–95° E) (b). (The annual cycle over the Atlantic is similar to that over the Pacific shown here, with slightly farther-southward (down to 2° N) excursions of the ITCZ in boreal winter.) The ITCZ (precipitation maxima) is marked by red lines. The seasonal ITCZ migration is sinusoidal with a moderate amplitude over the Pacific, away from continents; zonal winds remain easterly year-round (a). The seasonal ITCZ migration features abrupt and large shifts in the South Asian

monsoon sector, marking the onset and retreat of the summer monsoon; zonal winds north of the Equator turn westerly at monsoon onset (b, see Box 1). The precipitation data are the daily TMPA data⁶² averaged over 1998–2012. The data are smoothed temporally and meridionally by robust local linear regressions, spanning 11 days in time and 1° in latitude. The wind data are the 10-m winds from the ECMWF interim reanalysis⁴⁴ for the same years. The longest wind vector (in the South Asian monsoon sector at 18° S in September) corresponds to a wind speed of 9.1 m s^{-1} , and vector components to the left and right indicate westward and eastward wind components, respectively.

show that the ITCZ neither simply follows the insolation maximum^{5,16–18}, nor the sinusoidal seasonal variations of the interhemispheric temperature contrast. What mechanisms control the position of the ITCZ and its rainfall intensity is an important unanswered question in climate dynamics.

The ITCZ rainfall is fed by warm and moist trade winds near the surface (Figs 1a and 2). Their convergence leads to ascent of air masses, cooling, condensation and precipitation from deep convective clouds. In the upper troposphere, the air masses detrain from the clouds and diverge. They flow away from the ITCZ in the zonal mean, descend in the subtropics, and flow back along the surface towards the ITCZ, forming the meridionally

overturning Hadley circulation (Fig. 4). The ITCZ must be understood as one facet of meridional overturning circulations and zonal overturning circulations such as the Walker circulation: as the location of their ascending branches. The ITCZ varies with the overturning circulations. The atmospheric circulations, in turn, interact with circulations in the underlying oceans, which modify the thermal conditions at the surface that drive the atmospheric circulations (Fig. 4).

Observations and simulations indicate that the ITCZ migrates meridionally and its rainfall intensity changes when the atmospheric energy balance shifts^{19–32}. Elucidating how such ITCZ variations occur is the purpose

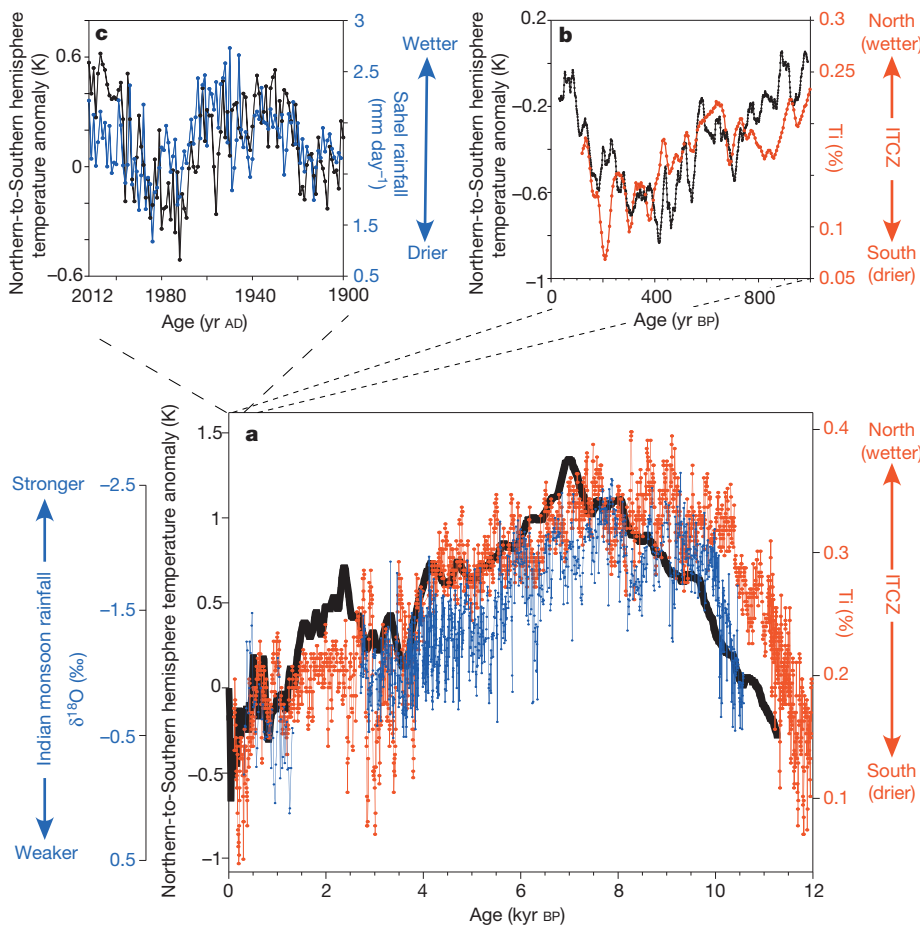


Figure 3 | Holocene ITCZ migrations, Indian monsoon and interhemispheric temperature contrast. a, Past 12,000 years. The temperature contrast between the extratropics of the Northern and Southern hemispheres (black) is strongly correlated with proxies of runoff into the Cariaco basin (red) and of Indian monsoon rainfall (blue). The temperature contrast is derived from the regularized expectation-maximization reconstructions in ref. 78, for areas poleward of 30° N and 30° S. The Cariaco runoff proxy is elemental titanium (Ti) in the ODP site 1002 sediment core (Fig. 1a), smoothed by a three-point running mean⁷. Higher Ti concentrations indicate more terrestrial runoff and rainfall, interpreted as a farther-norward ITCZ excursion in boreal summer^{6,7}. Declining Ti concentrations indicate a southward migration of the boreal-summer ITCZ over the later Holocene. The Indian monsoon proxy is the abundance of oxygen-18 ($\delta^{18}\text{O}$), relative to the Vienna Pee Dee Belemnite (VPDB) standard, in radiometrically dated cave stalagmites in Oman⁹ (Fig. 1a). Increasing $\delta^{18}\text{O}$ values indicate weakening summer monsoon rainfall. b, Past 1,000 years. Ti in Cariaco sediments (red), with a higher-resolution reconstruction⁹⁸ of Northern Hemisphere temperatures (black, 20-point running mean). (Southern Hemisphere temperatures changed little during this time⁸³.) c, Past century. Temperature contrast between Northern and Southern hemisphere extratropics (poleward of 24° N and 24° S) based on instrumental data⁶⁰ (black), and average daily rainfall over the Sahel (12°–18° N, 20° W–35° E) during June–October based on land station data⁹⁹ (blue). All temperatures and temperature contrasts are given as anomalies relative to the 1960–91 AD mean.

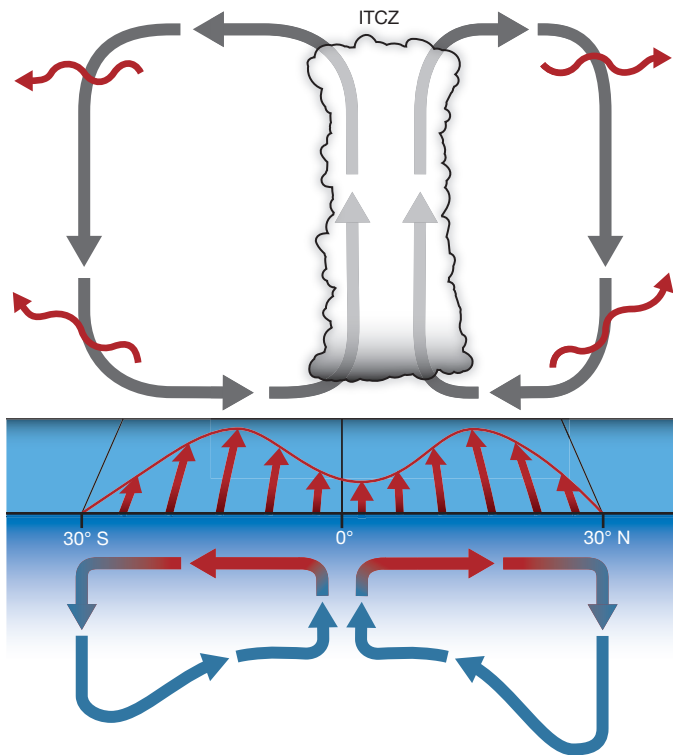


Figure 4 | Processes controlling zonal-mean ITCZ position. The lower branches of the Hadley circulation (grey arrows) bring warm and moist air masses towards the ITCZ, where they converge, rise and diverge as cooler and drier air masses aloft. Because the moist static energy aloft is greater than near the surface, the Hadley circulation transports energy away from the ITCZ. Eddies transport that energy farther into the extratropics (red wavy arrows). Hemispheric asymmetries in the energy export out of the tropics generally lead to an energy flux that crosses the Equator. Currently, the energy export into the extratropics in the south exceeds that in the north, leading to a southward cross-equatorial energy flux (Fig. 5). This implies an ITCZ in the Northern Hemisphere. Coupled to the Hadley circulation are mean zonal winds (red arrows at the sea surface), which are easterly where the near-surface mass flux is equatorward, and westerly where it is poleward. In the oceans, these zonal winds drive subtropical cells, with near-surface mass flux to the right of zonal winds in the Northern Hemisphere, and to the left in the Southern Hemisphere. Water masses cool and sink along their way towards the Hadley circulation termini and return below the sea surface (red and blue arrows). With mean easterlies in the tropics, the returning cool water masses upwell at the Equator, and the subtropical cells transport energy away from the Equator. But the upwelling location can migrate with the ITCZ away from the Equator and can dampen the ITCZ migration (Box 1).

of this Review. What emerges is a framework that links ITCZ variations to the energy input to and energy fluxes in the atmosphere. It allows us to interpret ITCZ variations across timescales from years to geological epochs.

Atmospheric energy balance and dynamics

Energy flux equator and ITCZ position

Although the air masses diverging in the upper troposphere above the ITCZ are cooler and drier than those converging near the surface, their potential energy is greater, such that their moist static energy—the energy relevant for transport considerations—is generally greater than that of the near-surface air masses^{33,34}. Therefore, vertically integrated over atmospheric columns, deep overturning circulations such as the Hadley circulation transport energy in the direction of their upper branches: away from the ITCZ (Fig. 1b). Averaged over a span of longitudes wide enough that one can focus on meridional fluxes, the ITCZ can be expected to lie near the “energy flux equator”^{21,22}, where the atmospheric meridional energy flux F changes sign—insofar as eddy contributions to the tropical atmospheric energy flux divergence remain negligible^{30,33}. Because the energy flux F usually increases going

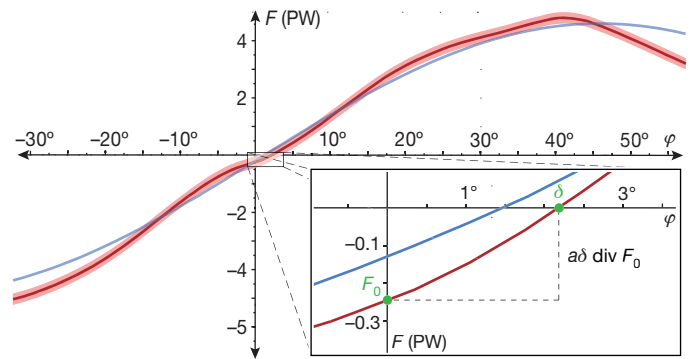


Figure 5 | Atmospheric meridional energy flux and energy flux equator. The atmospheric moist static energy flux F in the zonal and annual mean in the present climate (red line) is generally poleward, but it has a small southward component F_0 at the Equator. The energy flux equator is the zero of the energy flux, which currently lies around $\delta \approx 2.5^\circ$. Given the equatorial values of the energy flux F_0 and of its ‘slope’ with latitude $\text{div} F_0$, the energy flux equator δ can be determined from $F_0 \approx -a \text{div} F_0$, where a is Earth’s radius. For example, if F_0 increases (indicated schematically by the blue line), the energy flux equator δ moves southward. Similarly, if $\text{div} F_0$ increases, the energy flux equator moves towards the Equator. The energy flux data are from the ECMWF interim reanalysis for 1998–2012, corrected as in ref. 37 and provided by the National Center for Atmospheric Research. The light red shading indicates an estimated ± 0.2 PW standard error (the actual uncertainty is poorly known).

northward in Earth’s tropics—its divergence $\text{div} F$ is usually positive, meaning that energy is exported out of the tropics (Fig. 1b)—one expects the energy flux equator and the ITCZ to lie farther north the stronger southward is the cross-equatorial energy flux F_0 (Fig. 5). This is indeed what is seen in observations and climate simulations^{20–30,32}. Moreover, for a fixed cross-equatorial energy flux F_0 , one expects the energy flux equator and the ITCZ to lie closer to the Equator for a steeper equatorial ‘slope’ $\text{div} F$ of the energy flux with latitude (Fig. 5)³⁵.

More precisely, the atmospheric energy balance³³

$$\text{div} F = S - L - O \quad (1)$$

connects the divergence of the atmospheric energy flux F to the net energy input to the atmosphere, consisting of the net downward shortwave radiation S at the top of the atmosphere, minus the outgoing longwave radiation L and any ocean energy uptake O owing to transport or storage in the oceans. Energy storage on land is negligible on timescales of seasons and longer, as is storage in the atmosphere, at least in the tropics³⁶, on which we focus. Now we consider a zonal average over a span of longitudes (for example, an ocean basin) sufficiently wide that zonal fluxes can be ignored. By expanding the meridional energy flux F to first order in the latitude δ of the energy flux equator, we obtain $0 = F_\delta \approx F_0 + (\text{div} F_0)a\delta$, where the subscripts δ and 0 indicate latitude, and a is Earth’s radius (Fig. 5). Solving for the energy flux equator gives³⁵

$$\delta \approx -\frac{1}{a} \frac{F_0}{S_0 - L_0 - O_0} \quad (2)$$

Hence, the energy flux equator, and approximately the ITCZ position, depend to first order on the cross-equatorial atmospheric energy flux F_0 and on the net energy input to the atmosphere at or near the Equator: $\text{div} F_0 = S_0 - L_0 - O_0$. To be sure, the energy flux equator is not as sharply defined as is the ITCZ, because the nearly moist adiabatic thermal stratification implies that the atmospheric energy flux near the ITCZ is weak³⁴ (Fig. 1b). The energy flux equator also does not always coincide with the ITCZ (over the annual cycle^{5,26} for example). But its meridional excursions have magnitudes similar to those of the ITCZ^{5,22,35}, so equation (2) provides a starting point for understanding the ITCZ position quantitatively.

In the present climate in the zonal and annual mean, the atmosphere transports 0.3 ± 0.2 PW of energy southward across the Equator (Fig. 5)^{30,37}. The net equatorial energy input to the atmosphere³⁷ is $18 \pm 3 \text{ W m}^{-2}$ (Fig. 1b). With that, equation (2) implies an energy flux equator at $4^\circ \text{ N} \pm 3^\circ$ —broadly

consistent with the actual energy flux equator (Fig. 5) and the mean ITCZ position³⁵. However, the net equatorial energy input, $S_0 - L_0 - O_0$, is a small residual of large terms^{37,38}: $S_0 \approx 323 \text{ W m}^{-2}$, $L_0 \approx 251 \text{ W m}^{-2}$ and $O_0 \approx 54 \text{ W m}^{-2}$. Reducing S_0 or increasing L_0 or O_0 by only 6 W m^{-2} (by 2%–10%) suffices to move the energy flux equator δ a factor of about 1.5 farther poleward. The energy flux equator and ITCZ are sensitive to slight energetic shifts.

The mean position of the ITCZ in the Northern Hemisphere is thus linked to the atmospheric energy transport, which is directed from the warmer Northern Hemisphere into the 1.2–1.5 K cooler Southern Hemisphere^{30,31,39}. The Northern Hemisphere is warmer primarily because the Atlantic's meridional overturning circulation (AMOC) transports energy northward, up the mean temperature gradient. It dominates the cross-equatorial ocean energy transport; the resulting net northward transport across the Equator amounts to about 0.5 PW in the zonal mean^{30,37,39}. Some of this ocean energy transport across the Equator is compensated by the southward (downgradient) atmospheric energy transport, primarily accomplished by a Hadley cell with ascending branch and ITCZ north of the Equator (Fig. 4). (The Hadley cell, in turn, is often dominated by regional overturning cells, such as in the Asian monsoon sector¹⁸.) The AMOC energy transport displaces the ITCZ north of the Equator also in the Pacific (Fig. 2a), because winds homogenize the effect of AMOC energy transport zonally in the extratropics; hence, the partially compensating atmospheric energy transport is more zonally uniform and affects the ITCZ similarly over the Atlantic and Pacific³². This global energetic perspective de-emphasizes atmosphere–ocean interactions triggered by the shape of coastlines, which had previously been suggested to control the mean ITCZ position^{2,40}. Such local processes probably still play a role in shaping zonal asymmetries, but they appear to be secondary in displacing the Atlantic and Pacific ITCZ north of the Equator^{28,30,31}. However, local processes do seem to be responsible for the annual-mean ITCZ position south of the Equator over the Indian Ocean (Fig. 1a). The southern ITCZ arises because in the Indian Ocean a secondary precipitation maximum is maintained south of the Equator even in boreal summer (Fig. 2b), probably because the northward monsoonal flow rises and generates precipitation south of the Equator⁴¹, before crossing the Equator in the free troposphere and continuing towards the primary convergence zone farther north.

More generally, cross-equatorial atmospheric energy flux F_0 into one hemisphere demands an ITCZ in the opposite hemisphere, provided $\text{div } F_0 > 0$. Equation (2) quantifies how much the ITCZ migrates northward when F_0 decreases (becomes more southward), and how much the ITCZ migrates equatorward when $S_0 - L_0 - O_0$ increases. Studies hitherto have focused on the cross-equatorial energy flux—or closely related meridional temperature gradients (see below). But the energy flux alone cannot determine the ITCZ position, because it has units of power rather than units of ITCZ displacement. The net equatorial energy input matters too. Often the two will vary simultaneously, so that the regression coefficient of the ITCZ position on F_0 alone, inferred in recent studies^{24,26}, generally differs from the corresponding coefficient in a bivariate regression on F_0 and $S_0 - L_0 - O_0$ (ref. 35). How the ITCZ globally and regionally depends on both should be analysed in simulations and observations.

However, equation (2) is merely an energy balance identity that does not in itself imply a direction of causality. To establish causal links, it is necessary to discuss how energetic shifts arise mechanistically.

Mechanisms of triggering ITCZ migrations

Simulations and records of past climates have established that hemispherically asymmetric changes in the extratropics—for example, in AMOC energy transport or ice cover—can trigger ITCZ migrations⁵. To explain how extratropical changes can shift the energy flux equator and ITCZ, we need to discuss a special property of tropical temperatures^{21,23}.

Tropical temperatures above the planetary boundary layer are dynamically constrained to vary only weakly in the horizontal⁴² and to be nearly symmetric about the Equator. For example, in a Hadley circulation that is unaffected by eddy fluxes of angular momentum and that has an ITCZ far from the Equator, free-tropospheric temperatures are exactly symmetric

about the Equator, despite the strong asymmetry of the circulation⁴³. Indeed, hemispherically asymmetric variations of mid-tropospheric temperatures⁴⁴ currently do not exceed approximately 3 K in the zonal mean between 20° N and 20° S, not even in boreal summer, when strongly asymmetric monsoon circulations dominate the Hadley circulation.

The near-symmetry of tropical temperatures about the Equator means that a hemispherically asymmetric atmospheric energy export out of the tropics (say, across 30° N or 30° S) generally drives an energy flux across the Equator. To see this, assume that the tropics are confronted with a strengthening of the energy export across their northern boundary, which may be triggered by reduced AMOC energy transport or increased Arctic ice cover. First, we take cloud radiative effects and ocean energy uptake to be fixed, so that S and O stay fixed. In response to the perturbation at the northern boundary, temperatures may adjust throughout the tropics, but in the free troposphere they will remain nearly symmetric about the Equator—and so will the response of L , which depends on temperatures at the mid-tropospheric emission levels (for a fixed distribution of longwave absorbers and, in particular, clouds). Therefore, the response of $S - L - O$ is nearly symmetric about the Equator. The energy balance (equation (1)) then implies that the response of the energy flux F consists of a constant component F_0 and a component $F(\phi) \approx -F(-\phi)$ that is nearly antisymmetric about the Equator. The antisymmetric component is associated with hemispherically nearly symmetric energy export out of the tropics. So, to the extent that any hemispherically asymmetric energy export persists, it drives an energy flux F_0 through the tropics³⁵, leading to a northward cross-equatorial energy flux perturbation in our example. The energy flux equator (equation (2)), and with it the ITCZ, migrate southward.

Feedbacks from clouds, the distribution of water vapour, and ocean energy uptake modulate the ITCZ response, both by modulating $S_0 - L_0 - O_0$ and by inducing hemispheric asymmetries in the off-equatorial $S - L - O$ response, which modulate the F_0 response³⁵. But because the longwave and shortwave effects of the deep ITCZ clouds nearly cancel^{38,45}, their feedback on the net radiative energy input $S - L$ is small. Other tropical cloud feedbacks may modulate $S - L$ further²². Yet their overall effect appears to be small; even its sign is uncertain⁴⁶. The specific humidity of the atmosphere is enhanced near the ITCZ and reduces clear-sky L there. This, by itself, amplifies ITCZ migrations by increasing $S - L - O$ in the hemisphere with the ITCZ and amplifying the cross-equatorial energy flux into the opposite hemisphere. But because clear-sky variations in L near the ITCZ are only a small contributor to hemispheric asymmetries in L ($\leq 4 \text{ W m}^{-2}$ in the annual mean according to satellite data³⁸), this feedback is probably weak. Ocean energy uptake O may also cause hemispheric asymmetries in $S - L - O$. It and the shallow subtropical cells that dominate the ocean energy transport in low latitudes^{47–49} interact with the ITCZ. When the ITCZ migrates sufficiently far from the Equator, upwelling of cold waters and ocean energy uptake is enhanced near the ITCZ and dampens the ITCZ migration (Box 1). In the Indian Ocean, this feedback dampens the seasonal ITCZ migrations of the South Asian monsoon^{50,51}. It is probably even more important on longer timescales.

Thus, modulated by tropical feedbacks, hemispherically asymmetric energy export out of the tropics generally drives a cross-equatorial atmospheric energy flux. In the zonal and annual mean today, the atmosphere exports around 4 PW energy poleward across both 30° N and 30° S (Fig. 5), but it exports around 0.6 PW more southward than northward³¹. This in itself implies a southward atmospheric energy transport across the Equator of about 0.3 PW, consistent with the observed flux^{30,37} $F_0 \approx -0.3 \pm 0.2 \text{ PW}$. (A constant energy flux F_0 through the tropics adds to the energy export out of the tropics in one hemisphere and subtracts from it in the other, implying a hemispheric difference of $2F_0$ in energy exports.) Hemispheric asymmetries in $S - L - O$ within the tropics³⁵ contribute less to F_0 . The southward cross-equatorial energy flux is extratropically triggered in that AMOC energy transport differentially reduces meridional temperature gradients in the northern extratropics. This weakens the energy export out of the northern tropics because it weakens the poleward eddy energy flux³⁰, which dominates the energy export³³ and generally strengthens both with increasing temperature gradients and temperatures (because specific humidities

BOX 1

Atmosphere–ocean coupling and feedback on ITCZ

Ekman balance, $fV = -\tau^x$ (where the Coriolis parameter is f , and the eastward surface stress is τ^x), holds outside the equatorial latitude belt between approximately 5° N and 5° S and links zonal surface winds to ageostrophic near-surface meridional mass fluxes in the atmosphere and oceans. The meridional Ekman mass flux V is directed to the right of the zonal surface stress τ^x in the Northern Hemisphere ($f > 0$), and to its left in the Southern Hemisphere ($f < 0$).

In the atmosphere, Ekman balance accounts for the near-surface mean meridional mass flux. The stress τ^x retards zonal surface winds u and thus has sign opposite to that of u . An equatorward near-surface flow ($fV < 0$) implies surface easterlies ($u < 0$). This is the case over the tropical Atlantic and Pacific year-round: the ITCZ does not migrate sufficiently far from the Equator to lead to poleward near-surface flow off the Equator (Fig. 2a). A poleward near-surface flow ($fV > 0$) off the Equator implies surface westerlies ($u > 0$), just as over the tropical Indian Ocean during the summer monsoon: the ITCZ migrates so far off the Equator that the zonal winds turn westerly (Fig. 2b).

In the oceans, Ekman balance accounts for most of the near-surface mean meridional mass flux in the shallow subtropical cells^{47–49}. By Newton's third law, the wind stress τ^x driving V in the oceans is equal and opposite to the stress that retards the atmosphere's zonal winds over the oceans. Therefore, the near-surface Ekman mass flux is poleward ($fV > 0$) where zonal winds are easterly ($\tau^x < 0$), and equatorward otherwise. Remarkably, because the mean zonal stress on the atmosphere is equal and opposite to that on the oceans (up to the small surface stresses over land), the zonal-mean Ekman mass fluxes ($\pm \tau^x/f$) in the low-latitude atmosphere and oceans are approximately equal and opposite—even though water is a thousand times denser than air⁴⁹.

Thus the oceanic subtropical cells and the Hadley cells overhead resemble each other. Mirroring the upper branches of the Hadley cells, near-surface waters in the upper branches of the subtropical cells flow meridionally away from the ITCZ. They cool and are subducted on their way towards the Hadley circulation termini, where zonal surface winds turn westerly. The cooler water masses return below the surface and upwell near the ITCZ^{47,48} (Fig. 4). This ocean circulation transports energy away from the upwelling region. Its zonal- and annual-mean energy transport is currently about as strong as the atmospheric energy transport⁴⁹, but it weakens as the climate warms¹⁰⁰. Today, upwelling occurs at the Equator in the Atlantic and Pacific because mean zonal surface winds there are easterly year-round, so Ekman mass fluxes diverge at the Equator (f changes sign). But when the ITCZ is far enough from the Equator that tropical surface westerlies occur—such as in the Indian Ocean during the summer monsoon—upwelling occurs near the ITCZ, and the ocean can transport energy away from it. This increases the ocean energy uptake O and reduces the atmospheric energy input to the hemisphere with the ITCZ. It reduces the cross-equatorial atmospheric energy transport into the opposite hemisphere, thus dampening ITCZ migrations.

and latent energy fluxes increase with temperature^{35,52,53}. Ultimately, this chain of processes set in motion by AMOC energy transport leads to the mean position of the ITCZ north of the Equator^{30,31}. Conversely, if AMOC energy transport were weaker or absent, the atmospheric energy export out of the tropics to the north may well exceed that to the south: the more prominent continentality of the Northern Hemisphere leads to stronger stationary eddies, which enhance the poleward eddy energy flux³³. This may result in a northward F_0 and an ITCZ south of the Equator.

For the extratropical–tropical communication, then, it suffices to perturb the atmospheric energy export out of the tropics. Processes such as

wind-induced evaporation, previously proposed as mediators between the extratropics and tropics²⁰, are not essential^{25,32}. This energetic perspective is consistent with the notion that the ITCZ typically lies in the warmer hemisphere and migrates farther into that hemisphere the larger the inter-hemispheric temperature contrast^{19–32}. Because tropical temperatures are nearly symmetric about the Equator, the warmer hemisphere will usually have weaker meridional temperature gradients in the extratropics, implying reduced energy export out of the tropics^{52,53} and an ITCZ in that hemisphere. The larger the interhemispheric temperature contrast, the stronger is the unilateral reduction in energy export (other factors equal), and the farther in the warmer hemisphere is the ITCZ. Through closures for the atmospheric energy flux^{23,24}, the energy flux equator (equation (2)) and the ITCZ position have been quantitatively connected to temperatures, temperature gradients and other (for example, humidity) variables³⁵.

Towards a complete theory

The atmospheric energy balance is a starting point for understanding the ITCZ. But it gives an incomplete picture. The energy input to the atmosphere depends on the circulation, which must also satisfy other balances. In particular, the angular momentum balance is important for the Hadley circulation and convergence zones^{53–57}. For example, the Hadley circulation undergoes a regime transition from an equinox regime, in which its angular momentum balance is dominated by eddy fluxes, to a monsoonal solstice regime, in which its angular momentum balance is less affected by eddy fluxes^{18,55}. This regime transition is rendered abrupt by dynamical feedbacks within the Hadley circulation and in its interaction with extratropical eddies—feedbacks that may account for the square-wave shape of seasonal ITCZ migrations in the South Asian monsoon sector (Fig. 2b)¹⁸. In contrast, the energy flux equator migrates sinusoidally; it does not capture the ITCZ migrations fully⁵. For a full account, the energetic perspective must be paired with an account of the Hadley circulation—an outstanding challenge whose resolution may now be within reach because today's computational capabilities allow us to interrogate Hadley circulation dynamics systematically in experiments with global circulation models^{53,56}.

An account of the Hadley circulation is also necessary for a complete theory of how rainfall intensity in the ITCZ varies. Net precipitation (precipitation minus evaporation, measured, for example, by palaeo-proxies of continental runoff) is balanced by the convergence of water vapour in the atmosphere. Its variations consist of two components: a dynamic component associated with variations of the mass fluxes advecting the water vapour, and a thermodynamic component associated with variations of the specific humidity of the air masses¹². The dynamic component includes migrations of the ITCZ and variations of Hadley circulation strength. Like the ITCZ position, the Hadley circulation strength depends on extratropical temperature gradients, among other factors, and so will also respond to extratropical changes^{53,55}. Therefore, dynamic variations of ITCZ net precipitation can be expected to respond to Earth's orbital precession (which alters temperature gradients seasonally) and obliquity variations (which alter temperature gradients in the annual mean)^{12,14}, in addition to the response to variations in $S_0 - L_0 - O_0$. In contrast, thermodynamic variations of ITCZ net precipitation depend locally on the tropical energy balance, which controls the specific humidity⁵³. The tropical energy balance in the annual mean is not affected by precession and is only weakly affected by obliquity variations⁵⁸. But precession influences the specific humidity seasonally, and it can lead to an annual-mean thermodynamic rainfall response because seasonal humidity changes at a location are only relevant when the ITCZ is overhead, whereas annual-mean evaporation responses are weak⁵⁹. A rich interplay of such dynamic and thermodynamic responses to different orbital variations probably caused the tropical rainfall variations indicated by palaeo-records.

Observations of ITCZ variations

Direct observations in the past century and palaeo-records for the more distant past afford instructive case studies of how the ITCZ varies with climate and how it shifts with the atmospheric energy balance.

Direct observations in the past century

El Niño and Southern Oscillation (ENSO). The alternation between warm El Niño and cold La Niña conditions in the eastern Pacific provides a counterexample to the notion that the ITCZ generally migrates towards a differentially warming hemisphere. Going from typical La Niña to El Niño conditions, annual-mean temperatures rise by around 0.1 K averaged globally, but the extratropics of the Northern Hemisphere warm 0.08 K more than those of the Southern Hemisphere (according to data from ref. 60). Yet the ITCZ shifts southward⁶¹. The shift is clearest in boreal winter: precipitation data⁶² show that the western and central Pacific ITCZ shifts southward by about 2° going from typical La Niña to El Niño conditions, and by about 5° during strong El Niño events (such as in 1983 and 1998). Additionally, the longitudinal ITCZ structure is modified by ENSO's zonal rearrangement of convection⁶¹.

The energy balance provides an explanation. Zonal-mean O_0 in boreal winter is approximately 15 W m^{-2} smaller during El Niño than during La Niña, with the largest changes in the eastern Pacific (according to the data described in refs 37 and 44; see also ref. 63). This reduction in ocean energy uptake increases $S_0 - L_0 - O_0$, which in the zonal mean in boreal winter more than doubles from typical La Niña to El Niño. It implies that the displacement of the energy flux equator (equation (2)) and of the ITCZ away from the Equator is reduced by more than a factor of two—sufficient to account for the observed shift in ITCZ position. The effect of changes in F_0 is smaller (about a quarter of the total shift). Thus, ENSO variations of the ITCZ position appear to be primarily driven by tropical changes in the atmospheric energy input.

Differential hemispheric warming. Increasing concentrations of greenhouse gases have led to global warming over the past century, which, however, was not uniform²⁷. Until the 1930s, the northern extratropics warmed relative to the southern. Then, from the 1950s to the 1970s, the northern extratropics cooled by around 0.6 K relative to the southern extratropics. Since the mid-1970s, the northern extratropics have been warming again relative to the southern extratropics, thus far by around 0.7 K (Fig. 3c). (Differential warming rates for the entire hemispheres, rather than just the extratropics, are similar.) Differential warming of the northern extratropics is expected as concentrations of greenhouse gases increase: continents warm more than oceans and are more prominent in the Northern Hemisphere^{64–66}. The differential cooling of the northern extratropics in the mid-twentieth century has been ascribed to increasing concentrations of anthropogenic aerosols in the Northern Hemisphere⁶⁷, and to natural variations particularly in the North Atlantic^{68,69}.

To the extent that tropical temperatures are symmetric about the Equator, warming of the northern relative to the southern extratropics differentially reduces meridional temperature gradients in the northern extratropics. This by itself should lead to differentially reduced eddy energy export out of the tropics into the Northern Hemisphere and a northward ITCZ migration. The converse holds for warming of the southern relative to the northern extratropics.

Tropical precipitation records are not accurate enough to assess past ITCZ migrations in detail. But there is evidence that the zonal-mean ITCZ migrated southward as the northern extratropics cooled differentially in the mid-twentieth century^{70,71}; simulations show a similar response to increased anthropogenic aerosol loadings⁷⁰. Consistently, the northward ITCZ excursion in boreal summer appears to have been capped at lower latitudes, at least over the Atlantic⁷². The neighbouring semi-arid African Sahel region, which lies at the northern limit of the seasonal ITCZ excursion over the continent, dried severely (Fig. 3c)^{73–75}. Indeed, decadal variations of Sahel rainfall over the past century correlate with variations of the interhemispheric temperature contrast, including the differential warming of the northern extratropics (increasing Sahel rainfall) early in the twentieth century and since the 1980s (Fig. 3c).

Palaeo-records going back to the last ice age

Palaeoclimate data do not yet allow a global analysis of ITCZ variations but indicate globally coherent regional migrations. They suggest that the ITCZ position depends robustly on the interhemispheric temperature

contrast, which—since the last ice age—has varied primarily because of: (1) Earth's precessional cycle⁵⁸, with a period of about 21 kyr; and (2) millennial climate changes around the North Atlantic, driven by changes in AMOC and glaciation^{76,77}.

Clear evidence for ITCZ migrations comes from the Cariaco basin^{6,7} (Fig. 1a). There, the ITCZ during today's summers is overhead, trade winds and the ocean upwelling they drive are weak, and rainfall and runoff are at a maximum. Under such conditions, Cariaco sediments accumulate dark terrestrial detritus rich in elements such as titanium. In winter, the ITCZ migrates southward, trade winds and ocean upwelling strengthen, and runoff into the Cariaco basin is reduced. This increases the production of lightly coloured marine biogenic debris, which comes to dominate the flux to the sea bed⁶. From this alteration of Cariaco sediments, ITCZ migrations can be reconstructed.

Holocene. The Holocene (the past 11.7 kyr) represents the current interglacial (Fig. 3). Over its first 2–4 kyr, the Northern Hemisphere warmed⁷⁸ as glacial ice retreated⁷⁹. But after the Holocene thermal maximum at around 8 kyr BP, one of the few remaining drivers of climate variations was the precessional cycle. Early in the Holocene, perihelion occurred during boreal summer, intensifying summer insolation in the Northern Hemisphere and weakening it in the Southern Hemisphere. Perihelion then precessed towards its current occurrence in boreal winter. At the Holocene thermal maximum at around 8 kyr BP, the boreal-summer insolation contrast between the northern and southern extratropics was 9% stronger than it is today⁵⁸. The implied summertime warming of the northern extratropics relative to the southern should have led to a farther northward seasonal excursion of the ITCZ. (The excursion is modulated by equatorial insolation, which enters equation (2) through S_0 and was 5% stronger in boreal summer at 8 kyr BP; however, changes in S_0 are at least partially compensated by changes in L_0 and hence only represent a comparatively weak modulation of $S_0 - L_0 - O_0$.) Indeed, Cariaco sediments indicate greater summertime runoff around 8 kyr BP, followed by a gradual reduction until the beginning of the LIA at 500 BP (Fig. 3a, b). This signals a southward migration of the boreal-summer ITCZ, possibly paired with a thermodynamic reduction in rainfall intensity because of weakening local summer insolation. Cariaco runoff correlates well with reconstructions of the interhemispheric temperature contrast over the Holocene, which exhibits variations with amplitude ($\sim 0.6 \text{ K}$) similar to those seen over the past century (Fig. 3a,c)^{7,78}.

Consistent with the trend in the Cariaco record, the Atlantic ITCZ appears to have migrated southward over the Holocene⁸⁰. Tropical West Africa dried⁸¹ as the northern extratropics cooled differentially—the counterpart of the Sahel drying in the mid-twentieth century. The African Humid Period, during which lakes existed in the Sahara, appears to have terminated abruptly at around 5 kyr BP (ref. 81), perhaps when the boreal-summer ITCZ excursion over Africa ceased to reach the catchment areas of the sites from which the terminations are inferred. The complement to the reduction of boreal-summer rainfall in the northern tropics is a simultaneous increase of austral-summer rainfall in the southern tropics^{4,82}. This indicates a precession-driven farther-southward excursion of the ITCZ in austral summer, prompted by an increasing south-to-north interhemispheric insolation contrast (3% greater in austral summer today than 8 kyr BP). But precession-driven thermodynamic rainfall changes may also play a part⁵⁹.

Superimposed on the Holocene trend in the Cariaco record are abrupt features that cannot be explained by orbital insolation changes. The most notable is reduced runoff into the Cariaco basin during the LIA (Fig. 3b). The LIA cooling appears to have been most pronounced in high northern latitudes; it is absent in much of the Southern Hemisphere⁸³. Consistent with the implied differential cooling of the northern extratropics, Cariaco sediments suggest a southward ITCZ migration (Fig. 3b)⁷. Similarly, a variety of tropical Pacific records also indicate the ITCZ migrated southward during the LIA, followed by a northward migration at its end^{84,85}. Thus, the LIA appears to elicit a southward ITCZ migration similar to the more gradual differential Holocene cooling of the Northern Hemisphere.

Last ice age. North Atlantic climate records from the last ice age are rife with millennial-scale climate variations, such as the Dansgaard–Oeschger cycles, in which rapid warming is followed by gradual cooling that often

culminated in periods of massive iceberg discharge known as Heinrich stadials^{76,86}. Variations in the Southern Hemisphere are less pronounced, so changes in the interhemispheric temperature contrast were dominated by high northern latitudes⁸⁷. The origin of these cycles has eluded researchers for decades; AMOC changes are thought to play a role⁸⁸. Whatever their origin, the ITCZ appears to have responded to the millennial variations in high northern latitudes in the same fashion as to the more recent changes in interhemispheric temperature contrast. Each time high northern latitudes cooled, a dry interval is observed in the Cariaco basin, signalling a southward ITCZ migration (Fig. 6a, b). Consistent with a southward ITCZ migration, increased rainfall is indicated in the southern tropics^{4,89}.

Rainfall reductions coeval with strong high-latitude coolings are also indicated by cave stalagmites from Borneo (Fig. 1a), where today's ITCZ passes overhead in boreal spring and fall (Fig. 6c)^{90,91}. The fact that the Borneo record shows a response at least to the strongest millennial coolings in high latitudes suggests that it may likewise record ITCZ migrations. However, the slower orbital-timescale variations in the record correlate with local insolation when the ITCZ is overhead^{90,91}, suggesting that local thermodynamic rainfall variations may also play a part.

Monsoons and the ITCZ. Palaeo-records for the past 100 kyr indicate a strong correlation between the marine ITCZ position and monsoons¹³. Over the Holocene, as the boreal-summer position of the ITCZ migrated southward, Indian summer monsoon rainfall weakened⁹² (Fig. 3a). During the last ice age and the millennial cold intervals in high northern latitudes, Indian monsoon rainfall dropped, to be reinvigorated during subsequent warm intervals (Fig. 6d)¹¹. Cave stalagmites from eastern China^{92,93} and a host of other hydroclimate reconstructions⁹⁴ exhibit similar changes over the Holocene and millennial variations during the last ice age. They also correlate with precession-driven changes in northern high-latitude summer insolation⁹³, which seasonally influences extratropical temperature gradients and the energy export out of the tropics.

The similarity of marine ITCZ and monsoon variations may be taken as evidence that monsoon variations indicate rainfall redistribution within the

tropics and subtropics, with a northward displacement of the ITCZ implying increased rainfall farther north and decreased rainfall farther south^{12,13}. However, thermodynamic rainfall variations prompted by precession-driven local insolation variations may also have a role⁵⁹. The relative importance of dynamic and thermodynamic variations may shift with latitude and may account for the phase differences in rainfall variations seen, for example, between the Borneo and eastern China cave records^{90,91}.

Outlook

The energetic constraints on the ITCZ position have several implications. Stronger southward atmospheric energy transport across the Equator generally implies an ITCZ farther north; greater net energy input to the equatorial atmosphere implies an ITCZ closer to the Equator. The currently southward-directed cross-equatorial atmospheric energy transport weakens or reverses sign when the Northern Hemisphere extratropics cool relative to the southern extratropics, and this explains the southward ITCZ migrations that apparently occurred in the mid-twentieth century, LIA, and in cold intervals during the last ice age. Variations in the equatorial net energy input to the atmosphere modulate these ITCZ migrations, as they do during an El Niño. An outstanding challenge is to link orbital insolation variations to the shifts in the atmospheric energy balance that control the ITCZ position and its rainfall intensity—to arrive at a comprehensive understanding of tropical rainfall variations. Systematic studies with global climate models of how the tropical energy balance and extratropical energy fluxes respond to climate variations would be helpful. They will need to be accompanied by progress in our understanding of how clouds vary with climate and feed back onto the energy balance and ITCZ position.

The ITCZ position is sensitive to slight shifts in the atmospheric energy balance because the factors controlling it are small differences between large terms. The cross-equatorial energy flux depends on the small difference between the large energy exports out of the tropics into the northern and southern extratropics; the net energy input to the atmosphere is the small residual of the absorbed solar radiation, outgoing longwave radiation and

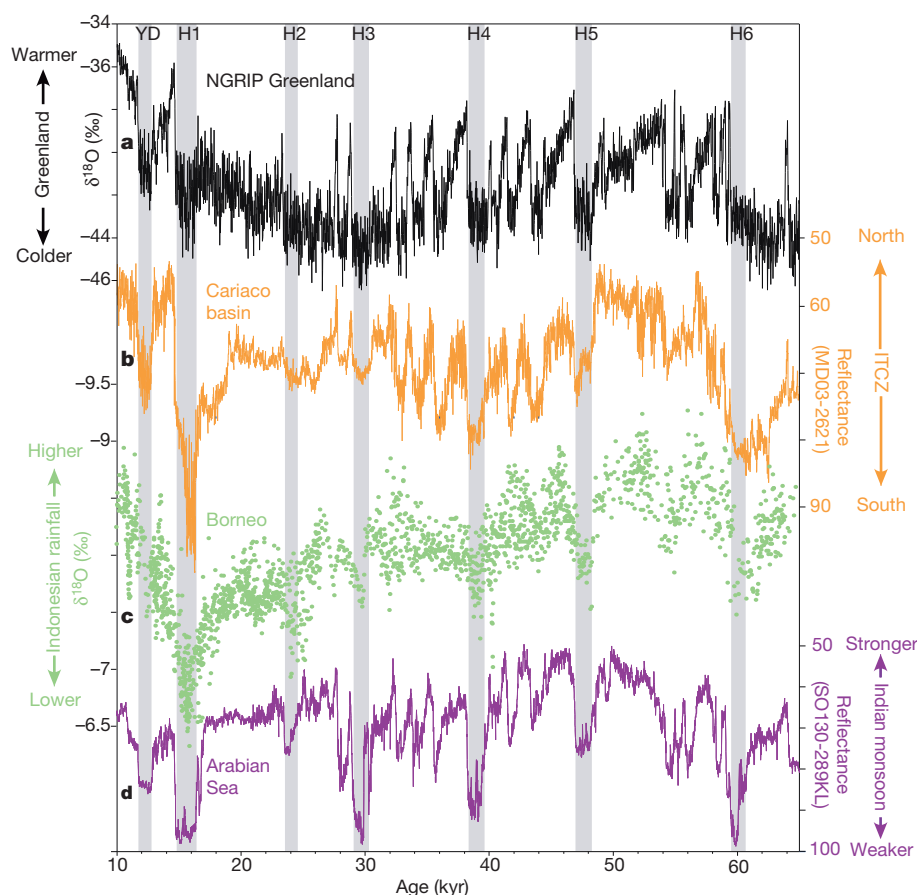


Figure 6 | Northern Hemisphere temperatures and ITCZ migrations during the last ice age. **a**, $\delta^{18}\text{O}$ from the North Greenland Ice Core Project (NGRIP) is a proxy of Arctic temperatures⁸⁶. It indicates gradual warming after the Last Glacial Maximum (~20 kyr BP) and millennial-scale Dansgaard-Oeschger cycles, between cold (stadial) and warmer (interstadial) intervals⁷⁶. Cold intervals include the Younger Dryas (YD) and Heinrich stadials H1 to H6 (grey shading). ($\delta^{18}\text{O}$ here is measured relative to Vienna Standard Mean Ocean Water (VSMOW).) **b**, Reflectance of Cariaco basin sediments measures the relative abundance of marine biogenic to terrigenous deposits. It is high when rainfall and runoff are low¹¹. Low reflectance is interpreted as a farther-northward ITCZ excursion in boreal summer. Warm intervals indicated by the Greenland ice core are generally associated with high runoff and a farther-northward boreal-summer ITCZ, and vice versa for cold intervals. During peak stadials, sediments lack the lamination produced by the annual rainfall cycle, indicating that the ITCZ appears to have been south of the Cariaco basin year-round¹¹. **c**, $\delta^{18}\text{O}$, relative to the VPDB standard, in cave stalagmites from Borneo is a proxy of rainfall in the equatorial western Pacific⁹¹. It is low during Heinrich stadials, but Dansgaard-Oeschger cycles otherwise are less evident than in the higher-latitude records. **d**, Analogously to **b**, low reflectance of Arabian Sea sediments indicates high runoff from Indian monsoon rainfall¹¹. It occurs during warm intervals indicated by the Greenland ice core. High reflectance and weaker Indian monsoon rainfall occur during cold intervals.

ocean energy uptake. Slight shifts in the large terms can lead to substantial ITCZ migrations. This sensitivity of the ITCZ position probably accounts for the difficulties climate models have in simulating it⁹⁵. Biases in simulating the ITCZ can arise, for example, through relatively small errors in representing clouds: in the extratropics²⁹, they distort the energy export out of the tropics, and in the tropics²³, they distort the net energy input. Yet this sensitivity also makes the ITCZ an excellent recorder of how energetic balances shift with climate. Reconstructions of past ITCZ migrations may soon constrain predictions about the future in regions such as the Sahel, where slight modifications of seasonal ITCZ excursions can change the hydroclimate drastically⁶⁶. As our theories advance and ITCZ and palaeo-reconstructions improve, it may become possible to use inferences about how past ITCZ migrations depend on insolation variations to constrain uncertain cloud feedbacks. Particularly helpful would be reconstructions of the ITCZ in the mid-Pliocene (4.5–3.1 million years BP), the youngest geological analogue of a climate that is 2–3 K warmer. The Northern Hemisphere then was ice-free and relatively warm, suggesting a more northern ITCZ, while the equatorial eastern Pacific had a weaker cold tongue^{96,97}, suggesting reduced ocean energy uptake and an ITCZ closer to the Equator. Additionally, net energy input to the atmosphere was modified by increased concentrations of greenhouse gases⁹⁶ and cloud changes. Knowing how these competing effects balanced would reduce uncertainties in predictions for the next century.

Received 25 November 2013; accepted 1 July 2014.

1. Waliser, D. E. & Gautier, C. A satellite-derived climatology of the ITCZ. *J. Clim.* **6**, 2162–2174 (1993).
2. Philander, S. *et al.* Why the ITCZ is mostly north of the equator. *J. Clim.* **9**, 2958–2972 (1996).
3. Gadgil, S. The Indian monsoon and its variability. *Annu. Rev. Earth Planet. Sci.* **31**, 429–467 (2003).
4. Koutavas, A. & Lynch-Stieglitz, J. in *The Hadley Circulation: Present, Past, and Future* (eds Diaz, H. F. & Bradley, R. S.) Vol. 21 *Advances in Global Change Research* 347–369 (Kluwer Academic, 2004).
5. Chiang, J. C. H. & Friedman, A. R. Extratropical cooling, interhemispheric thermal gradients, and tropical climate change. *Annu. Rev. Earth Planet. Sci.* **40**, 383–412 (2012).
- This paper reviews evidence for the dependence of the ITCZ position on the interhemispheric temperature contrast.**
6. Peterson, L. C., Haug, G. H., Hughen, K. A. & Rohl, U. Rapid changes in the hydrologic cycle of the tropical Atlantic during the last glacial. *Science* **290**, 1947–1951 (2000).
7. Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C. & Röhl, U. Southward migration of the Intertropical Convergence Zone through the Holocene. *Science* **293**, 1304–1308 (2001).
8. Koutavas, A., deMenocal, P. B., Olive, G. C. & Lynch-Stieglitz, J. Mid-Holocene El Niño–Southern Oscillation (ENSO) attenuation revealed by individual foraminifera in eastern tropical Pacific sediments. *Geology* **34**, 993–996 (2006).
9. Fleitmann, D. *et al.* Holocene forcing of the Indian monsoon recorded in a stalagmite from southern Oman. *Science* **300**, 1737–1739 (2003).
10. Fleitmann, D. *et al.* Holocene ITCZ and Indian monsoon dynamics recorded in stalagmites from Oman and Yemen (Socotra). *Quat. Sci. Rev.* **26**, 170–188 (2007).
11. Deplazes, G. *et al.* Links between tropical rainfall and North Atlantic climate during the last glacial period. *Nature Geosci.* **6**, 213–217 (2013).
12. Clement, A. C., Hall, A. & Broccoli, A. J. The importance of precessional signals in the tropical climate. *Clim. Dyn.* **22**, 327–341 (2004).
13. Cheng, H., Sinha, A., Wang, X., Cruz, F. W. & Edwards, R. L. The global paleomonsoon as seen through speleothem records from Asia and the Americas. *Clim. Dyn.* **39**, 1045–1062 (2012).
14. Prell, W. L. & Kutzbach, J. E. Sensitivity of the Indian monsoon to forcing parameters and implications for its evolution. *Nature* **360**, 647–652 (1992).
15. Wanner, H. *et al.* Mid- to Late Holocene climate change: an overview. *Quat. Sci. Rev.* **27**, 1791–1828 (2008).
16. Chou, C. & Neelin, J. D. Mechanisms limiting the northward extent of the northern summer monsoons over North America, Asia, and Africa. *J. Clim.* **16**, 406–425 (2003).
17. Xian, P. & Miller, R. L. Abrupt seasonal migration of the ITCZ into the summer hemisphere. *J. Atmos. Sci.* **65**, 1878–1895 (2008).
18. Bordoni, S. & Schneider, T. Monsoons as eddy-mediated regime transitions of the tropical overturning circulation. *Nature Geosci.* **1**, 515–519 (2008).
19. Vellinga, M. & Wood, R. A. Global climatic impacts of a collapse of the Atlantic thermohaline circulation. *Clim. Change* **54**, 251–267 (2002).
20. Chiang, J. C. H. & Bitz, C. M. Influence of high latitude ice cover on the marine Intertropical Convergence Zone. *Clim. Dyn.* **25**, 477–496 (2005).
21. Broccoli, A. J., Dahl, K. A. & Stouffer, R. J. Response of the ITCZ to northern hemisphere cooling. *Geophys. Res. Lett.* **33**, L01702 (2006).
22. Kang, S. M., Held, I. M., Frierson, D. M. W. & Zhao, M. The response of the ITCZ to extratropical thermal forcing: idealized slab-ocean experiments with a GCM. *J. Clim.* **21**, 3521–3532 (2008).
23. Kang, S. M., Frierson, D. M. W. & Held, I. M. The tropical response to extratropical thermal forcing in an idealized GCM: The importance of radiative feedbacks and convective parameterization. *J. Atmos. Sci.* **66**, 2812–2827 (2009).
- This paper imposes cross-equatorial ocean energy transport in an idealized general circulation model and, following ref. 21, interprets ITCZ response through the atmosphere–ocean energy balance.**
24. Frierson, D. M. W. & Hwang, Y.-T. Extratropical influence on ITCZ shifts in slab ocean simulations of global warming. *J. Clim.* **25**, 720–733 (2012).
25. Kang, S. M. & Held, I. M. Tropical precipitation, SSTs and the surface energy budget: a zonally symmetric perspective. *Clim. Dyn.* **38**, 1917–1924 (2012).
26. Donohoe, A., Marshall, J., Ferreira, D. & McGee, D. The relationship between ITCZ location and cross-equatorial atmospheric heat transport: from the seasonal cycle to the last glacial maximum. *J. Clim.* **26**, 3597–3618 (2013).
27. Friedman, A. R., Hwang, Y.-T., Chiang, J. C. H. & Frierson, D. M. W. Interhemispheric temperature asymmetry over the twentieth century and in future projections. *J. Clim.* **26**, 5419–5433 (2013).
28. Fuckar, N. S., Xie, S.-P., Farneti, R., Maroon, E. A. & Frierson, D. M. W. Influence of the extratropical ocean circulation on the Intertropical Convergence Zone in an idealized coupled general circulation model. *J. Clim.* **26**, 4612–4629 (2013).
29. Hwang, Y.-T. & Frierson, D. M. W. Link between the double-Intertropical Convergence Zone problem and cloud biases over the Southern Ocean. *Proc. Natl Acad. Sci. USA* **110**, 4935–4940 (2013).
30. Marshall, J., Donohoe, A., Ferreira, D. & McGee, D. The ocean's role in setting the mean position of the Inter-Tropical Convergence Zone. *Clim. Dyn.* **42**, 1967–1979 (2014).
- This paper shows that the northward AMOC energy transport is primarily responsible for the mean position of the ITCZ north of the Equator.**
31. Frierson, D. M. W. *et al.* Contribution of ocean overturning circulation to tropical rainfall peak in the northern hemisphere. *Nature Geosci.* **6**, 940–944 (2013).
32. Kang, S. M., Held, I. M. & Xie, S.-P. Contrasting the tropical responses to zonally asymmetric extratropical and tropical thermal forcing. *Clim. Dyn.* **42**, 2033–2043 (2013).
33. Peixoto, J. P. & Oort, A. H. *Physics of Climate* (American Institute of Physics, 1992).
34. Neelin, J. D. & Held, I. M. Modeling tropical convergence based on the moist static energy budget. *Mon. Weath. Rev.* **115**, 3–12 (1987).
35. Bischoff, T. & Schneider, T. Energetic constraints on the position of the Intertropical Convergence Zone. *J. Clim.* **27**, 4937–4951 (2014).
36. Donohoe, A. & Battisti, D. S. The seasonal cycle of atmospheric heating and temperature. *J. Clim.* **26**, 4962–4980 (2013).
37. Fasullo, J. T. & Trenberth, K. E. The annual cycle of the energy budget. Part II: Meridional structures and poleward transports. *J. Clim.* **21**, 2313–2325 (2008).
38. Loeb, N. G. *et al.* Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Clim.* **22**, 748–766 (2009).
39. Feulner, G., Rahmstorf, S., Levermann, A. & Volkwardt, S. On the origin of the surface air temperature difference between the hemispheres in Earth's present-day climate. *J. Clim.* **26**, 7136–7150 (2013).
40. Xie, S.-P. in *The Hadley Circulation: Present, Past and Future* (eds Diaz, H. F. & Bradley, R. S.) Vol. 21 *Advances in Global Change Research* 121–152 (Kluwer Academic, 2004).
- This paper reviews how the shape of continents and atmosphere–ocean interactions can cause zonal asymmetries and displace the ITCZ north of the Equator.**
41. Pauluis, O. Boundary layer dynamics and cross-equatorial Hadley circulation. *J. Atmos. Sci.* **61**, 1161–1173 (2004).
42. Charney, J. G. A note on large-scale motions in the tropics. *J. Atmos. Sci.* **20**, 607–609 (1963).
43. Lindzen, R. S. & Hou, A. Y. Hadley circulations for zonally averaged heating centered off the equator. *J. Atmos. Sci.* **45**, 2416–2427 (1988).
44. Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).
45. Harrison, E. F. *et al.* Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment. *J. Geophys. Res.* **95**, 18687–18703 (1990).
46. Voigt, A., Stevens, B., Bader, J. & Mauritsen, T. Compensation of hemispheric albedo asymmetries by shifts of the ITCZ and tropical clouds. *J. Clim.* **27**, 1029–1045 (2014).
47. Schott, F. A., McCreary, J. P. Jr & Johnson, G. C. in *Earth's Climate: The Ocean–Atmosphere Interaction* Geophysical Monograph Series 147, 261–304 (American Geophysical Union, 2004).
48. Klinger, B. A. & Marotzke, J. Meridional heat transport by the subtropical cell. *J. Phys. Oceanogr.* **30**, 696–705 (2000).
49. Held, I. M. The partitioning of the poleward energy transport between the tropical ocean and atmosphere. *J. Atmos. Sci.* **58**, 943–948 (2001).
50. Jayne, S. R. & Marotzke, J. The dynamics of ocean heat transport variability. *Rev. Geophys.* **39**, 385–411 (2001).
51. Webster, P. J. in *The Asian Monsoon* (ed. Wang, B.) 3–66 (Springer Praxis, 2006).
52. Caballero, R. & Langen, P. L. The dynamic range of poleward energy transport in an atmospheric general circulation model. *Geophys. Res. Lett.* **32**, L02705 (2005).
53. Schneider, T., O'Gorman, P. A. & Levine, X. J. Water vapor and the dynamics of climate changes. *Rev. Geophys.* **48**, RG3001 (2010).
- This paper reviews how the hydrological cycle, the Hadley circulation, and extratropical energy transports vary with climate.**
54. Held, I. M. & Hou, A. Y. Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.* **37**, 515–533 (1980).

55. Walker, C. C. & Schneider, T. Eddy influences on Hadley circulations: simulations with an idealized GCM. *J. Atmos. Sci.* **63**, 3333–3350 (2006).
56. Schneider, T. The general circulation of the atmosphere. *Annu. Rev. Earth Planet. Sci.* **34**, 655–688 (2006).
57. Schneider, T. & Bordoni, S. Eddy-mediated regime transitions in the seasonal cycle of a Hadley circulation and implications for monsoon dynamics. *J. Atmos. Sci.* **65**, 915–934 (2008).
58. Berger, A. & Loutre, M. F. Insolation values for the climate of the last 10 million years. *Quat. Sci. Rev.* **10**, 297–317 (1991).
59. Merlis, T. M., Schneider, T., Bordoni, S. & Eisenman, I. The tropical precipitation response to orbital precession. *J. Clim.* **26**, 2010–2021 (2013).
60. Hansen, J., Ruedy, R., Sato, M. & Lo, K. Global surface temperature change. *Rev. Geophys.* **48**, RG4004 (2010).
61. Dai, A. & Wigley, T. M. L. Global patterns of ENSO-induced precipitation. *Geophys. Res. Lett.* **27**, 1283–1286 (2000).
62. Liu, Z., Ostrenga, D., Teng, W. & Kempler, S. Tropical Rainfall Measuring Mission (TRMM) precipitation data and services for research and applications. *Bull. Am. Meteorol. Soc.* **93**, 1317–1325 (2012).
63. Trenberth, K. E., Caron, J. M., Stepaniak, D. P. & Worley, S. Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures. *J. Geophys. Res.* **107**, 4065, <http://dx.doi.org/10.1029/2000JD000298> (2002).
64. Joshi, M. M., Gregory, J. M., Webb, M. J., Sexton, D. M. H. & Johns, T. C. Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Clim. Dyn.* **30**, 455–465 (2008).
65. Byrne, M. P. & O’Gorman, P. A. Link between land–ocean warming contrast and surface relative humidities in simulations with coupled climate models. *Geophys. Res. Lett.* **40**, 5223–5227 (2013).
66. Broecker, W. S. & Putnam, A. E. Hydrologic impacts of past shifts of Earth’s thermal equator offer insight into those to be produced by fossil fuel CO₂. *Proc. Natl Acad. Sci. USA* **110**, 16710–16715 (2013).
67. Tett, S. F. B. *et al.* Estimation of natural and anthropogenic contributions to twentieth century temperature change. *J. Geophys. Res.* **107**, 4306 <http://dx.doi.org/10.1029/2000JD000028> (2002).
68. Schneider, T. & Held, I. M. Discriminants of twentieth-century changes in Earth surface temperatures. *J. Clim.* **14**, 249–254 (2001).
69. Thompson, D. W. J., Wallace, J. M., Kennedy, J. J. & Jones, P. D. An abrupt drop in northern hemisphere sea surface temperature around 1970. *Nature* **467**, 444–447 (2010).
70. Rotstayn, L. D. & Lohmann, U. Tropical rainfall trends and the indirect aerosol effect. *J. Clim.* **15**, 2103–2116 (2002).
- This paper demonstrates that increased atmospheric aerosol loading can lead to southward ITCZ migrations and presents evidence that this occurred in the twentieth century.**
71. Hwang, Y.-T., Frierson, D. M. W. & Kang, S. M. Anthropogenic sulfate aerosol and the southward shift of tropical precipitation in the late 20th century. *Geophys. Res. Lett.* **40**, 2845–2850 (2013).
72. Tokinaga, H. & Xie, S.-P. Weakening of the equatorial Atlantic cold tongue over the past six decades. *Nature Geosci.* **4**, 222–226 (2011).
73. Folland, C. K., Palmer, T. N. & Parker, D. E. Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature* **320**, 602–607 (1986).
74. Giannini, A., Saravanan, R. & Chang, P. Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science* **302**, 1027–1030 (2003).
75. Held, I. M., Delworth, T. L., Lu, J., Findell, K. L. & Knutson, T. R. Simulation of Sahel drought in the 20th and 21st centuries. *Proc. Natl Acad. Sci. USA* **102**, 17891–17896 (2005).
- This paper presents an overview of the mid-20th century Sahel drought, including its causes and predictions for the future.**
76. Bond, G. *et al.* Correlations between climate records from North Atlantic sediments and Greenland ice. *Nature* **365**, 143–147 (1993).
77. Broecker, W. Massive iceberg discharges as triggers for global climate change. *Nature* **372**, 421–424 (1994).
78. Marcott, S. A., Shakun, J. D., Clark, P. U. & Mix, A. C. A reconstruction of regional and global temperature for the past 11,300 years. *Science* **339**, 1198–1201 (2013).
79. Renssen, H. *et al.* The spatial and temporal complexity of the Holocene thermal maximum. *Nature Geosci.* **2**, 411–414 (2009).
80. Arbuszewski, J. A., deMenocal, P. B., Cléroux, C., Bradtmiller, L. & Mix, A. Meridional shifts of the Atlantic intertropical convergence zone since the Last Glacial Maximum. *Nature Geosci.* **6**, 959–962 (2013).
- This paper infers ITCZ migrations from reconstructions of ocean salinity, which diversifies evidence for ITCZ migrations and may eventually allow reconstructions of rainfall at the ITCZ.**
81. deMenocal, P. *et al.* Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing. *Quat. Sci. Rev.* **19**, 347–361 (2000).
82. Baker, P. A. *et al.* Tropical climate changes at millennial and orbital timescales on the Bolivian Altiplano. *Nature* **409**, 698–701 (2001).
83. Jones, P. D., Briffa, K. R., Barnett, T. P. & Tett, S. F. B. High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with general circulation model control-run temperatures. *Holocene* **8**, 455–471 (1998).
84. Linsley, B. K., Dunbar, R. B., Wellington, G. M. & Mucciarone, D. A. A coral-based reconstruction of Intertropical Convergence Zone variability over Central America since 1707. *J. Geophys. Res.* **99**, 9977–9994 (1994).
85. Sachs, J. P. *et al.* Southward movement of the Pacific Intertropical Convergence Zone AD 1400–1850. *Nature Geosci.* **2**, 519–525 (2009).
- This paper presents evidence that the ITCZ across the Pacific was farther south during the Little Ice Age.**
86. Wolff, E. W., Chappellaz, J., Blunier, T., Rasmussen, S. O. & Svensson, A. Millennial-scale variability during the last glacial: The ice core record. *Quat. Sci. Rev.* **29**, 2828–2838 (2010).
87. Blunier, T. *et al.* Asynchrony of Antarctic and Greenland climate change during the last glacial period. *Nature* **394**, 739–743 (1998).
88. Broecker, W. S. Thermohaline circulation, the Achilles heel of our climate system: Will man-made CO₂ upset the current balance? *Science* **278**, 1582–1588 (1997).
89. Wang, X. *et al.* Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. *Nature* **432**, 740–743 (2004).
90. Partin, J. W., Cobb, K. M., Adkins, J. F., Clark, B. & Fernandez, D. P. Millennial-scale trends in west Pacific warm pool hydrology since the Last Glacial Maximum. *Nature* **449**, 452–455 (2007).
91. Carolin, S. A. *et al.* Varied response of western Pacific hydrology to climate forcings over the last glacial period. *Science* **340**, 1564–1566 (2013).
92. Wang, Y. *et al.* The Holocene Asian monsoon: links to solar changes and North Atlantic climate. *Science* **308**, 854–857 (2005).
- This paper shows that Asian monsoon rainfall weakened over the past 8 kyr.**
93. Wang, Y. *et al.* Millennial and orbital-scale changes in the East Asian monsoon over the past 224,000 years. *Nature* **451**, 1090–1093 (2008).
94. An, Z. S. The history and variability of the East Asian paleomonsoon climate. *Quat. Sci. Rev.* **19**, 171–187 (2000).
95. Lin, J.-L. The double-ITCZ problem in IPCC AR4 coupled GCMs: ocean–atmosphere feedback analysis. *J. Clim.* **20**, 4497–4525 (2007).
96. Fedorov, A. V. *et al.* Patterns and mechanisms of early Pliocene warmth. *Nature* **496**, 43–49 (2013).
97. Zhang, Y. G., Pagani, M. & Liu, Z. A 12-million-year temperature history of the tropical Pacific Ocean. *Science* **344**, 84–87 (2014).
98. Moberg, A., Sonechkin, D. M., Holmgren, K., Datsenko, N. M. & Karlén, W. Highly variable northern hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature* **433**, 613–617 (2005).
99. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
100. Levine, X. J. & Schneider, T. Response of the Hadley circulation to climate change in an aquaplanet GCM coupled to a simple representation of ocean heat transport. *J. Atmos. Sci.* **68**, 769–783 (2011).

Acknowledgements J. Fasullo and K. Trenberth from the National Center for Atmospheric Research provided the energy flux data we used in Figs 1b and 5 and in some of the estimates in the text. The top-of-atmosphere radiative flux estimates in the text are based on NASA Clouds and the Earth’s Radiant Energy System (CERES) data, version CERES EBAF-TOA Ed2.7. S. Marcott provided the temperature reconstructions in Fig. 3, and M. Hell drew Figs 4 and 5. We are grateful for discussions with D. Sigman and N. Meckler and for comments on drafts by F. Ait-Chaalal, A. Donohoe, R. Ferrari, and J.-E. Lee. The research underlying this paper was supported by grants from the US National Science Foundation (numbers AGS-1019211, AGS-1049201 and AGS-1003614).

Author Contributions All authors discussed the central concepts and ideas and processed and analysed data. T.S. and G.H.H. led the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence should be addressed to T.S. (tapio@ethz.ch).

Developmental plasticity and the origin of tetrapods

Emily M. Standen¹, Trina Y. Du² & Hans C. E. Larsson²

The origin of tetrapods from their fish antecedents, approximately 400 million years ago, was coupled with the origin of terrestrial locomotion and the evolution of supporting limbs. *Polypterus* is a member of the basal-most group of ray-finned fish (actinopterygians) and has many plesiomorphic morphologies that are comparable to elpistostegid fishes, which are stem tetrapods. *Polypterus* therefore serves as an extant analogue of stem tetrapods, allowing us to examine how developmental plasticity affects the ‘terrestrialization’ of fish. We measured the developmental plasticity of anatomical and biomechanical responses in *Polypterus* reared on land. Here we show the remarkable correspondence between the environmentally induced phenotypes of terrestrialized *Polypterus* and the ancient anatomical changes in stem tetrapods, and we provide insight into stem tetrapod behavioural evolution. Our results raise the possibility that environmentally induced developmental plasticity facilitated the origin of the terrestrial traits that led to tetrapods.

The evolution of terrestrial locomotion in vertebrates required the appearance of new behaviours and supporting appendicular structures^{1–8}. The skeletal changes included the origin of supporting limbs, the decoupling of the dermal pectoral girdle from the skull and the strengthening of the girdle ventrally for support⁹. The predicted behavioural changes at this transition include the planting of the pectoral fins closer to the midline of the body, thereby increasing the vertical component of the ground reaction force and raising the anterior body off the ground^{4,5,10}. How these evolutionary changes arose at the origin of tetrapods is still largely unclear, and the evolutionary processes surrounding these ancient events are not accessible because the transitional animals are extinct.

Environmentally induced phenotypes and their subsequent incorporation into heritable material may play an important role in macroevolution, including in the origin of novel traits^{11,12}. Phenotypic plasticity is the ability of an organism to react to the environment by changing its morphology, behaviour, physiology and biochemistry¹¹. Such responses are often beneficial for the survival and fitness of an organism and may facilitate success in novel environments^{13–15}. Phenotypically plastic traits can also eventually become heritable through genetic assimilation, which fixes a reduced range of phenotypic plasticity by decreasing a trait’s environmental sensitivity^{11,16,17}.

Major transitions that involved the modification and appearance of complex or novel traits may be accessible through the existing developmental pathways that allow plasticity in an ancestor. The ‘flexible stem’ model describes a process by which the fixed phenotypes of ecologically specialized lineages reflect the assimilation of alternative phenotypes by the ancestral lineage^{11,18–20}. For example, when marine sticklebacks were raised on alternative diets, developmental plasticity in the head shape and mouth shape paralleled the phenotypic divergence in the derived ecotypes¹⁹. Examining plasticity in an extant form may therefore shed light on the epigenetic processes in past evolutionary events^{11,12,19}.

This Article relates plasticity in an extant fish taxon to a major evolutionary transition: the origin of tetrapods. The plasticity of ancient fish might have provided the variation necessary to allow the evolution of the terrestrially functional fins that eventually evolved into limbs. Validating such a prediction is difficult as stem tetrapods are extinct. In these cases,

a sister taxon to the derived groups of interest can be used to estimate the ancestral plasticity¹².

In this study, we investigated developmental plasticity in *Polypterus*, the extant fish closest to the common ancestor of actinopterygians and sarcopterygians²¹. We chose *Polypterus* because it is one of the best models for examining the role of developmental plasticity during the evolution of stem tetrapods. *Polypterus* has an elongate body form, rhomboid scales, ventrolaterally positioned pectoral fins and functional lungs, all traits that are comparable to elpistostegid fishes. Moreover, this living fish is capable of surviving on land and can perform tetrapod-like terrestrial locomotion with its pectoral fins (E.M.S., T.Y.D., P. Laroche, B. Wilhelm and H.C.E.L., manuscript in preparation). Other terrestrially locomotory fish are derived teleosts that use a range of different gaits and have derived morphologies. Extant sarcopterygians (that is, lungfish (Dipnoi) and coelacanths (*Latimeria chalumnae*)) do not use their fins for terrestrial locomotion. By using an animal that already displays a walking behaviour, such as *Polypterus*, we can compare how obligatory walking influences gait and skeletal structure.

We raised control and treatment groups of *Polypterus* under aquatic and terrestrial conditions, respectively, to examine their behavioural and anatomical plasticity in response to a terrestrial habitat. By placing this predominantly aquatic animal in an obligatory terrestrial environment, we changed the forces experienced by the animal’s musculo-skeletal system. We predicted that the increased gravitational and frictional forces experienced by terrestrialized fish would cause changes in the ‘effectiveness’ of their locomotory behaviour when travelling over land, as well as changes in the shape of the skeletal structures used in locomotion. We also predicted that the plastic responses of the pectoral girdle of terrestrialized *Polypterus* would be similar to the directions of the anatomical changes seen in the stem tetrapod fossil record.

Swimming versus walking behaviour

During steady swimming, *Polypterus* oscillates its pectoral fins for propulsion, with little body and tail motion. *Polypterus* can also walk on land using a contralateral gait, by using its pectoral fins to raise its head and anterior trunk off the ground and by using its posterior body for forward

¹Biology Department, University of Ottawa, Gendron Hall, 30 Marie Curie Private, Ottawa, Ontario K1N 6N5, Canada. ²Redpath Museum, McGill University, 859 Sherbrooke Street West, Montreal, Quebec H3A 0C4, Canada.

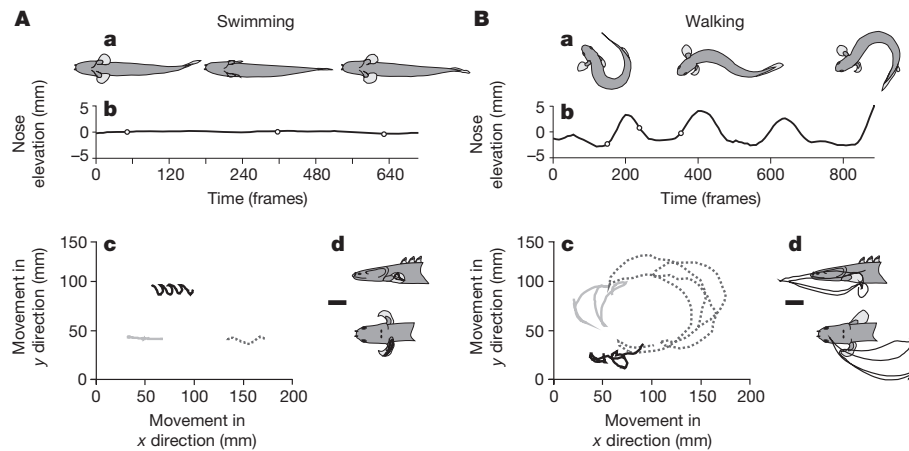


Figure 1 | Kinematic behaviour of swimming and walking *Polypterus*.

A, Swimming in one exemplar fish. **B**, Walking in one exemplar fish. **a**, Maximum and minimum body curvature over one stroke cycle. **b**, Change in nose elevation over several stroke cycles (filmed at 250 frames s^{-1}). The circles correspond to the illustrations (from left to right) in **a**. **c**, Movement in the x - y plane (parallel to the ground) of the nose (solid grey line), tail (dotted grey line)

propulsion (Supplementary Video 1; E.M.S., T.Y.D., P. Laroche, B. Wilhelm and H.C.E.L., manuscript in preparation). We observed critical performance differences between swimming and walking (Fig. 1 and Extended Data Table 1). When *Polypterus* swam with a pectoral fin gait, it moved farther and faster per fin beat than when it walked. When walking, the fish moved their bodies and fins faster, and their nose, tail and fin oscillations were larger. Walking fish also had higher nose elevations, longer stroke durations and greater body curvatures. These performance differences suggest that walking is energetically more expensive than swimming²².

Biomechanical response to living on land

The treatment group of fish (which was raised on land) walked differently from the control group of fish (which was raised in water) (Extended Data Tables 1 and 2). The fish that were raised on land had their fins planted on the ground for less time and had shorter stride durations. Despite the faster stroke cycle of the treatment group fish, both treatment and control fish had comparable duty factors (approximately 0.65). Planted fins always slipped during the fin plant phase of the stroke, but the duration and distance of this potentially expensive fin 'slip' were shorter in the fish raised on land. The land-raised fish also had smaller pectoral fin excursions and elevations, planted their fins closer to the body midline, and had higher nose elevations and smaller tail oscillations. Reduction of unnecessary fin and tail motion and fine-tuning of fin placement to reduce body friction with the ground are required to minimize energy expenditure during walking, suggesting that the terrestrially raised *Polypterus* has a more efficient gait.

The timing of critical kinematic variables is also important when determining performance. During walking, for both groups of fish, the fin was first planted near its most forward position, and the body was maximally curved, with the tail closest to the head. With the fin grounded, the tail pushed forwards, vaulting the anterior body and head over the planted fin. This positioning suggests that both fin and tail work together as the main thrust producers. The body curvature, tail amplitude, nose elevation and the start of the fin plant had similar timings in the stroke cycle for both groups of fish despite the differences in the environments in which the fish were raised (Fig. 2a, Extended Data Table 3 and Supplementary Video 1). These similarities indicate that walking with fins has basic kinematic requirements that are not necessarily improved with practise or competence. By contrast, the treatment group fish had a predictable timing for the majority of walking variables, whereas the control fish exhibited high degrees of variation. Importantly, the fin slide start occurred later during walking in the land-raised fish than in the water-raised fish. Additionally, the maximum forward (abduction) and

backward (adduction) fin position, the end of the fin plant, the end of the fin slide and the minimum distance from the fin to the body midline occurred at specific times within the stroke cycle in land-raised fish but were unpredictable in water-raised fish (Fig. 2 and Extended Data Table 3). These differences in timing between the control and treatment fish groups may indicate that the fish raised on land have a conditioned 'training' advantage.

During walking, in all fish, the maximum nose elevation and the beginning of the fin slide occurred half way through the head swing, when the nose crossed the forward path (Fig. 2b). The co-occurrence of these factors supports the hypothesis that the fin acts as a pole over which the head and anterior body are vaulted. Ground friction caused by the body and the amount of time the head and body are held off the ground must be limited and balanced to reduce the energy expenditure during walking. The treatment group fish had higher nose elevations and shorter

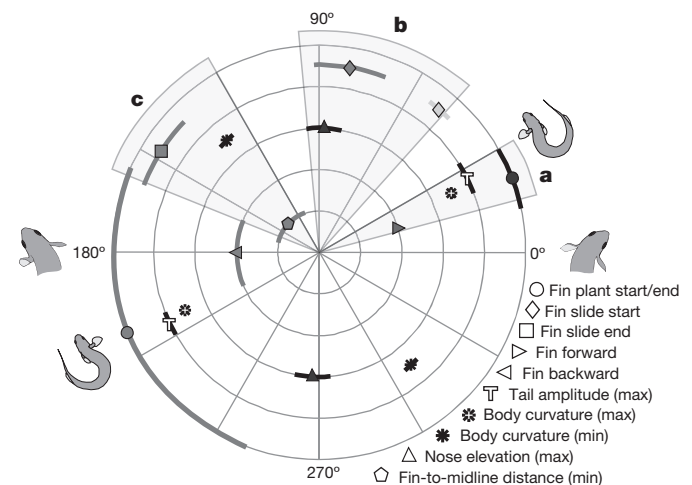


Figure 2 | Timing of kinematic variables for the left fin during walking in control and treatment group fish. One complete stroke cycle is represented by 360°. The stroke starts at 0° (head on the right). Mid-stroke occurs at 180° (head on the left). The stroke ends at 360° (0°). Data with significant directionality (Rayleigh's test, $P < 0.05$) are plotted for the treatment group (dark grey, $n = 12$) and the control group (light grey, $n = 6$). Variables with similar timing between groups ($P > 0.05$) were binned and plotted as one (black). The symbols represent the mean timing (\pm angular variance, shown as a line) of the different kinematic variables. Some symbols occlude an extremely small angular variance. The areas highlighted in light grey (**a**, **b**, **c**) are key kinematic points in the stroke cycle. max, maximum; min, minimum.

stroke cycles, suggesting that the body ground friction and head elevation time were minimized. Additionally, the treatment group fish had less fin and body motion, less fin slip and less time between the start of the fin slide and the maximum head elevation, with the head and anterior body more effectively vaulting over the fin, possibly minimizing the energy loss due to slip. This difference in what seems to be a fundamentally important characteristic in walking over land hints that differences in control and/or 'effectiveness' during walking depend on an individual's training environment.

The treatment group fish also had a minimized distance between their planted fin and their body midline at the same time their fin stopped sliding (Fig. 2c). Bringing the fin closer to the midline generates more vertical ground reaction force through the fin, explaining the reduced fin slip. Control fish planted their fins farther from their body midline. These variables are critical for minimizing the effort required to move the body weight over land²³. By precisely controlling fin placement and fin slide timing, the treatment group fish may be streamlining the power that is required by the tail and the body to push the fish forwards by ensuring that the tail and body thrust occur when the body and head are lifted by the fin. Furthermore, the treatment group fish kept their fins stationed on the ground for the remainder of the step, allowing the fin to contribute to force production and control during the final phase of the step, and for the initiation of the next contralateral step. These behavioural differences are hypothesized to be 'learned' training advantages. Conversely, control fish had more variable fin slip timing, which lasted longer, as well as inconsistent timing patterns for the end of the fin plant, suggesting that they had not optimized their biomechanical performance.

The minimal differences in the kinematic variables between the treatment and control groups during swimming indicate that there was minimal 'loss' of swimming function associated with being raised in a terrestrial environment without the ability to 'practise' swimming after gill absorption (Extended Data Tables 1 and 4, Extended Data Fig. 1 and Methods).

Anatomical response to living on land

Like the biomechanical properties, the pectoral anatomy of land-raised *Polypterus* also exhibited phenotypic plasticity in response to terrestriation. The clavicle, cleithrum and supracleithrum of the fish pectoral girdle create a supporting brace that links the head and the body during locomotion and feeding (Fig. 3a). In most fish, including *Polypterus*, the paired clavicles are ventral, attaching to the anterior tip of the cleithrum medially and joining in symphysis at the midline, acting as a structural support for the neck and pectoral girdle²⁴. The cleithrum serves as an attachment point for muscles connecting to the skull anteriorly, to the trunk muscles posteriorly and to the pectoral fin muscles laterally²⁵. The supracleithrum articulates with the cleithrum and joins the pectoral girdle to the skull via the posttemporal bone.

The clavicle and cleithrum had significantly different shapes in the land-raised and water-raised groups (Fig. 3). The treatment group fish had narrower and more elongated clavicles, with more pointed processes that were 10.6% longer ($P \leq 0.046$). The clavicle cross-section was thinner in treatment group fish, in which the clavicle forms a cup that conforms to the cleithrum contact (Fig. 3c). In the land-raised fish, the cleithrum's horizontal arm also had a narrower lateral surface (Fig. 3b). Increased forces in the pectoral girdle from gravitational and postural changes may have induced a modelling response of the clavicle and cleithrum²⁶. Differences in bone shape may also reflect the need for increased fin mobility in terrestrial environments. When *Polypterus* walks, its fins must move through a larger range of motion than when it swims, forcing the operculum to bend out of the way to accommodate forward fin excursion. The posterior margin of the opercular cavity is displaced caudally by a reduction of the cleithrum's lateral postbranchial lamina. This change, along with the increased length and slenderness of the clavicle in treatment group fish, expands the opercular cavity between the fin and the operculum, providing more space for the pectoral fins to move.

Overall, the supracleithrum did not have a significantly different shape in water-raised and land-raised fish, but it was the only bone with an

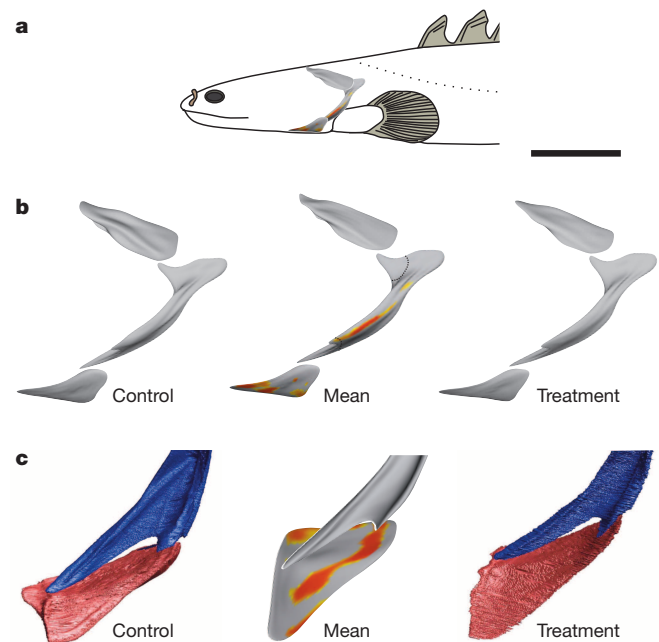


Figure 3 | Anatomical plasticity of *Polypterus* pectoral girdles. a, Location of the supracleithrum, cleithrum and clavicle in *Polypterus*. Scale bar, 1 cm. b, Left lateral views of the pectoral girdle with the mean clavicle (bottom), mean cleithrum (centre) and mean supracleithrum (top) dissociated for control (left) and treatment (right) group fish. A point-based multivariate analysis of covariance (MANCOVA) with correction for multiple comparisons (false discovery rate estimation) was used to determine the significant differences between the control ($n = 7$) and treatment ($n = 15$) *Polypterus* groups; significantly different regions are shown in colour (yellow, $P < 0.05$; red, $P < 0.01$). The mean illustration represents the average shape calculated from all individuals. c, Close-up anterolateral views of the mean clavicle (pink)–cleithrum (blue) contact in control (left) and treatment (right) group fish and the mean shape calculated from all individuals (centre).

allometric difference between the groups (Extended Data Table 5). The supracleithrum of the treatment group fish maintained the slenderness present at smaller bone sizes, with both the anterior process and posterior margin remaining narrow in the largest specimens and maintaining an underdeveloped midlateral ridge. The supracleithrum of control fish showed the opposite trend, developing robust anterior and posterior processes and midlateral ridges with increasing bone size. A size-related reduction in the supracleithrum's robustness may reflect a weakened connection to the posttemporal bone anteriorly and to the cleithrum posteriorly. This possibility suggests a weakened attachment between the pectoral girdle and the skull in land-raised fish. The size independence of the low midlateral ridge in the treatment group fish may also allow greater flexibility between the supracleithrum and the overlying operculum, further enlarging the opercular cavity and further freeing the pectoral girdle from the skull. Finally, comparison of bootstrap confidence intervals for shape variance showed that the land-reared fish had a higher shape disparity of the cleithrum and supracleithrum (Extended Data Fig. 2).

Plasticity and the origin of tetrapods

Terrestrialized *Polypterus* displayed less-variable walking behaviour, planted their pectoral fins closer to their body midlines, lifted their heads higher and had less fin slip, allowing more effective vaulting of the anterior body over the planted fin. These features are performance-enhancing traits during terrestrial locomotion¹⁰. We predict that similar behavioural changes were present in stem tetrapods. These locomotory changes in *Polypterus* probably affected the forces experienced by the skeleton, influencing skeletal growth and changing bone shape²⁷. The differences in bone morphology observed between terrestrialized and water-raised

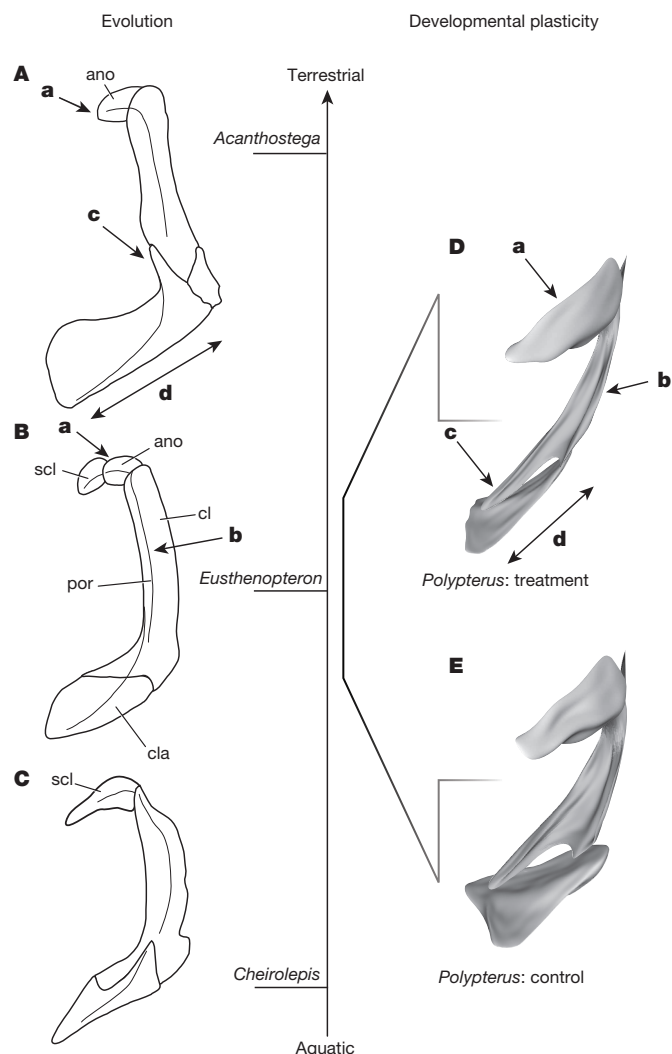


Figure 4 | Scenario for the contribution of developmental plasticity to large-scale evolutionary change in stem tetrapods. Left anterodorsolateral views of the pectoral girdle of selected stem tetrapods (A, B), an outgroup (C), and land-reared (D) and water-reared (E) *Polypterus*. The following are comparable developmentally plastic morphologies: reduction of the supracleithrum (a), reduction of the posterior opercular chamber edge (b), strengthened clavicle–cleithrum contact (c) and narrowing and elongation of the clavicle (d). ano, anocleithrum; cl, cleithrum; cla, clavicle; por, post opercular ridge (note the ridge in *Cheirolepis* is not distinct but is laterally positioned, as is shown); scl, supracleithrum.

Polypterus bear a remarkable resemblance to the evolutionary changes of stem tetrapod pectoral girdles during the Devonian period (Fig. 4). The skeletal changes seen in the treatment group fish revealed a marked reduction in the external boundaries of the opercular cavity bounded by the supracleithrum and cleithrum, which presumably facilitates greater flexibility between the pectoral girdle and the operculum, similar to what is observed in stem tetrapods such as *Eusthenopteron*²⁸. The elongation of the clavicles and the more tightly interlocking cleithrum–clavicle contact may strengthen the ventral brace through the clavicle, aiding in feeding, locomotion and body support in a terrestrial environment. Similar morphologies involving the medial bracing of the clavicles via an interclavicle are also thought to have stabilized the girdle in the earliest tetrapods, *Acanthostega* and *Ichthyostega*⁷. Finally, the dissociation of the pectoral girdle from the skull by reduction and loss of the supracleithrum and extrascapular bones allowed the evolution of a neck, an important feature for feeding on land⁹.

Novel or stressful environments, particularly those to which organisms have had no previous exposure or adaptations are catalysts for releasing variation^{29–31}. Evidence at a microevolutionary scale demonstrates that plasticity plays an important role in the appearance of complex traits^{18–20}. Our results show that exposure to a novel terrestrial environment can increase the phenotypic variation in the terrestrial locomotory behaviours and the pectoral girdle of *Polypterus*. We hypothesize that phenotypic plasticity, as a response to rapid and sustained environmental stresses, may also facilitate macroevolutionary change. Multi-generational experiments on terrestrialized *Polypterus* are required to determine the effect of developmental plasticity on the evolution of traits associated with effective terrestrial locomotion.

Developmental plasticity can be integrated into the study of major evolutionary transitions. The rapid, developmentally plastic response of the skeleton and behaviour of *Polypterus* to a terrestrial environment, and the similarity of this response to skeletal evolution in stem tetrapods, is consistent with plasticity contributing to large-scale evolutionary change. Similar developmental plasticity in Devonian sarcopterygian fish in response to terrestrial environments may have facilitated the evolution of terrestrial traits during the rise of tetrapods.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 March; accepted 24 July 2014.

Published online 27 August; corrected online 3 September 2014 (see full-text HTML version for details).

- Niedzwiedzki, G., Szrek, P., Narkiewicz, K., Narkiewicz, M. & Ahlberg, P. E. Tetrapod trackways from the early Middle Devonian period of Poland. *Nature* **463**, 43–48 (2010).
- Clack, J. A. The fin to limb transition: new data, interpretations, and hypotheses from paleontology and developmental biology. *Annu. Rev. Earth Planet. Sci.* **37**, 163–179 (2009).
- Callier, V., Clack, J. A. & Ahlberg, P. E. Contrasting developmental trajectories in the earliest known tetrapod forelimbs. *Science* **324**, 364–367 (2009).
- Shubin, N. H., Daeschler, E. B. & Jenkins, F. A. J. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature* **440**, 764–771 (2006).
- Pierce, S. E., Clack, J. A. & Hutchinson, J. R. Three-dimensional limb joint mobility in the early tetrapod *Ichthyostega*. *Nature* **486**, 523–527 (2012).
- Ahlberg, P. E. & Clack, J. A. Palaeontology: a firm step from water to land. *Nature* **440**, 747–749 (2006).
- Clack, J. A. *Gaining Ground* 2nd edn (Indiana Univ. Press, 2012).
- Coates, M. I. *Acanthostega gunnari* Jarvik: postcranial anatomy, basal tetrapod interrelationships and patterns of skeletal evolution. *Trans. R. Soc. Edinb. Earth Sci.* **87**, 363–421 (1996).
- Coates, M. I., Ruta, M. & Friedman, M. Ever since Owen: changing perspectives on the early evolution of tetrapods. *Annu. Rev. Ecol. Evol. Syst.* **39**, 571–592 (2008).
- Blob, R. W. & Biewener, A. A. Mechanics of limb bone loading during terrestrial locomotion in the green iguana (*Iguana iguana*) and American alligator (*Alligator mississippiensis*). *J. Exp. Biol.* **204**, 1099–1122 (2001).
- West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, 2003).
- Moczek, A. P. et al. The role of developmental plasticity in evolutionary innovation. *Proc. R. Soc. B* **278**, 2705–2713 (2011).
- Ghalambor, C. K., McKay, J. K., Carroll, S. P. & Reznick, D. N. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.* **21**, 394–407 (2007).
- Price, T. D., Qvarnstrom, A. & Irwin, D. E. The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. Lond. B* **270**, 1433–1440 (2003).
- Yeh, P. J. & Price, T. D. Adaptive phenotypic plasticity and the successful colonization of a novel environment. *Am. Nat.* **164**, 531–542 (2004).
- Crispo, E. The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* **61**, 2469–2479 (2007).
- Waddington, C. H. Genetic assimilation of an acquired character. *Evolution* **7**, 118–126 (1953).
- Gomez-Mestre, I. & Buchholz, D. R. Developmental plasticity mirrors differences among taxa in spadefoot toads linking plasticity and diversity. *Proc. Natl Acad. Sci. USA* **103**, 19021–19026 (2006).
- Wund, M. A., Baker, J. A., Clancy, B., Golub, J. L. & Foster, S. A. A test of the “flexible stem” model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am. Nat.* **172**, 449–462 (2008).
- Rajakumar, R. et al. Ancestral developmental potential facilitates parallel evolution in ants. *Science* **335**, 79–82 (2012).
- Inoue, J. G., Miya, M., Tsukamoto, K. & Nishida, M. Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the “ancient fish”. *Mol. Phylogenet. Evol.* **26**, 110–120 (2003).

22. Tucker, V. A. The energetic cost of moving about: walking and running are extremely inefficient forms of locomotion. Much greater efficiency is achieved by birds, fish—and bicyclists. *Am. Sci.* **63**, 413–419 (1975).
23. Kawano, S. M. & Blob, R. W. Propulsive forces of mudskipper fins and salamander limbs during terrestrial locomotion: Implications for the invasion of land. *Integr. Comp. Biol.* **53**, 283–294 (2013).
24. Gosline, W. A. The structure and function of the dermal pectoral girdle in bony fishes with particular reference to ostariophysines. *J. Zool.* **183**, 329–338 (1977).
25. Winterbottom, R. A descriptive synonymy of the striated muscles of the teleostei. *Proc. Acad. Nat. Sci. Philadelphia* **125**, 225–317 (1973).
26. Biewener, A. A. Scaling body support in mammals: limb posture and muscle mechanics. *Science* **245**, 45–48 (1989).
27. Frost, H. M. On our age-related bone loss: insights from a new paradigm. *J. Bone Miner. Res.* **12**, 1539–1546 (1997).
28. Andrews, S. M. & Westoll, T. S. The postcranial skeleton of *Eusthenopteron fori* Whiteaves. *Trans. R. Soc. Edinb.* **68**, 207–329 (1970).
29. Badyaev, A. V. Stress-induced variation in evolution: from behavioural plasticity to genetic assimilation. *Proc. R. Soc. B* **272**, 877–886 (2005).
30. Suzuki, Y. & Nijhout, H. F. Evolution of a polyphenism by genetic accommodation. *Science* **311**, 650–652 (2006).
31. Lerouzig, A. & Carlborg, O. Evolutionary potential of hidden genetic variation. *Trends Ecol. Evol.* **23**, 33–37 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank F. A. Jenkins Jr and C. R. Marshall for encouragement and discussions with E.M.S. during the initial phases of this project. We also thank E. Abouheif for discussions and editing of the manuscript. We thank J. Dawson for the loan of camera equipment, S. Bertram for access to laboratory space while running the experiment and B. Bongfeldt for animal care and preliminary data analysis. We are grateful for the Tomlinson Post-doctoral Fellowship (E.M.S.), for grants from the National Sciences and Engineering Research Council of Canada (PDF to E.M.S., CGS-M to T.Y.D. and Discovery Grant #261796-2011 to H.C.E.L.), for the Robert G. Goelet Research Award (E.M.S.) and to Canada Research Chairs (H.C.E.L.).

Author Contributions E.M.S. conceived, designed and conducted the experiments and biomechanical analyses. T.Y.D. scanned, segmented and analysed the micro-computed tomography images. H.C.E.L. provided palaeontological and evolutionary expertise that shaped the project. E.M.S., T.Y.D. and H.C.E.L. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.M.S. (estanden@uottawa.ca) or H.C.E.L. (hans.ce.larsson@mcgill.ca).

Alterations of the human gut microbiome in liver cirrhosis

Nan Qin^{1,2*}, Fengling Yang^{1*}, Ang Li^{1*}, Edi Prifti^{3*}, Yanfei Chen^{1*}, Li Shao^{1,2*}, Jing Guo¹, Emmanuelle Le Chatelier³, Jian Yao^{1,2}, Lingjiao Wu¹, Jiawei Zhou¹, Shujun Ni¹, Lin Liu¹, Nicolas Pons³, Jean Michel Batto³, Sean P. Kennedy³, Pierre Leonard³, Chunhui Yuan¹, Wenchao Ding¹, Yuanting Chen¹, Xinjun Hu¹, Beiwen Zheng^{1,2}, Guirong Qian¹, Wei Xu¹, S. Dusko Ehrlich^{3,4}, Shusen Zheng^{2,5} & Lanjuan Li^{1,2}

Liver cirrhosis occurs as a consequence of many chronic liver diseases that are prevalent worldwide. Here we characterize the gut microbiome in liver cirrhosis by comparing 98 patients and 83 healthy control individuals. We build a reference gene set for the cohort containing 2.69 million genes, 36.1% of which are novel. Quantitative metagenomics reveals 75,245 genes that differ in abundance between the patients and healthy individuals (false discovery rate < 0.0001) and can be grouped into 66 clusters representing cognate bacterial species; 28 are enriched in patients and 38 in control individuals. Most (54%) of the patient-enriched, taxonomically assigned species are of buccal origin, suggesting an invasion of the gut from the mouth in liver cirrhosis. Biomarkers specific to liver cirrhosis at gene and function levels are revealed by a comparison with those for type 2 diabetes and inflammatory bowel disease. On the basis of only 15 biomarkers, a highly accurate patient discrimination index is created and validated on an independent cohort. Thus microbiota-targeted biomarkers may be a powerful tool for diagnosis of different diseases.

Cirrhosis is an advanced liver disease resulting from acute or chronic liver injury, including alcohol abuse, obesity and hepatitis virus infection. The prognosis for patients with decompensated liver cirrhosis is poor, and they frequently require liver transplantation¹. The liver interacts directly with the gut through the hepatic portal and bile secretion² systems. Enteric dysbiosis, especially the translocation of bacteria³ and their products^{4,5} across the gut epithelial barrier, is involved in the progression of liver cirrhosis. However, the phylogenetic and functional composition changes in the human gut microbiota that are related to this progression remain obscure⁵. Some studies have revealed that alterations in the gut microbiota are important in complications of end-stage liver cirrhosis⁶ (such as spontaneous bacterial peritonitis⁷ and hepatic encephalopathy⁸) and the induction and promotion of liver damage in early-stage liver disease⁹ (such as alcoholic liver disease¹⁰ and non-alcoholic fatty liver disease¹¹), but definitive associations of gut microbiota and liver pathology in humans are still lacking¹². Studies of patients with liver cirrhosis¹³ and of mouse models for alcoholic liver disease¹⁰ have revealed a similar and substantial alteration in the gut microbiota, as measured by sequencing of 16S ribosomal RNA genes. How these phylogenetic alterations relate to changes in the functioning of this ecosystem is, however, unclear.

The role of gut microbiota in human health and disease¹⁴ has recently received considerable attention. Chronic diseases, such as obesity^{15–18}, inflammatory bowel disease (IBD)^{19,20}, diabetes mellitus^{21,22}, metabolic syndrome²³, symptomatic atherosclerosis²⁴ and non-alcoholic fatty liver disease¹⁰, have been associated with gut microbiota. The US National Institutes of Health Human Microbiome Project (HMP) generated a large data set from different anatomical sites among 242 healthy individuals and created a large human microbiome gene resource^{25,26}. Quantitative metagenomics analysis^{27,28} developed by the MetaHIT consortium revealed a significant loss of gut microbial richness associated with the

risk of metabolic syndrome related co-morbidities. Here we apply a similar analysis to contrast microbiota from 123 patients with liver cirrhosis and 114 healthy counterparts of Han Chinese origin.

Gene catalogue of gut microbes

We constructed a gene catalogue from 98 Chinese patients with liver cirrhosis and 83 healthy Chinese control individuals (Supplementary Table 1) using the methodology developed by MetaHIT. The liver cirrhosis catalogue contained 2,688,468 non-redundant open reading frames (ORFs). We compared it with three other gut microbial catalogues: MetaHIT²⁹, HMP²⁵ and T2D²². To facilitate this comparison, genes were predicted from the original contigs using the same criteria. The MetaHIT catalogue contained 3,452,726 genes, HMP 4,768,112 genes and T2D 2,148,029 genes. In total 674,131 genes were common to all catalogues (Extended Data Fig. 1a). The liver cirrhosis catalogue, MetaHIT, HMP and T2D gene sets contained 794,647, 1,419,517, 2,620,096 and 623,570 unique genes, respectively. Genes from the liver cirrhosis, T2D and MetaHIT catalogues were merged; the HMP was not included, as it contained Sanger, 454 or Illumina-based 16S sequences, in addition to whole metagenomic data. The merged non-redundant catalogue contained 5,382,817 genes (Extended Data Fig. 1b).

Phylogenetic profiles of gut microbes

The sequencing reads (36.67%) were aligned against 4,398 reference genomes from the National Center for Biotechnology Information and the HMP (Supplementary Table 2). After correction for population stratification that could be related to non-liver cirrhosis-related factors (see Methods), the relative abundances of phylum, class, order, family, genus and species between liver cirrhosis and control groups were compared (Extended Data Fig. 2). Phylotypes with a median relative abundance larger than 0.01% of the total abundance in either the healthy control

¹State Key Laboratory for Diagnosis and Treatment of Infectious Disease, The First Affiliated Hospital, College of Medicine, Zhejiang University, 310003 Hangzhou, China. ²Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Zhejiang University, 310003 Hangzhou, China. ³Metagenopolis, Institut National de la Recherche Agronomique, 78350 Jouy en Josas, France. ⁴King's College London, Centre for Host-Microbiome Interactions, Dental Institute Central Office, Guy's Hospital, London Bridge, London SE1 9RT, UK. ⁵Key Laboratory of Combined Multi-organ Transplantation, Ministry of Public Health, the First Affiliated Hospital, Zhejiang University, 310003 Hangzhou, China.

*These authors contributed equally to this work.

group or the liver cirrhosis group were included for comparison. At the phylum level, Bacteroidetes and Firmicutes dominated the faecal microbial communities of both groups (Fig. 1a, b). Compared with healthy controls, patients with liver cirrhosis had fewer Bacteroidetes (Fig. 1a), but higher levels of Proteobacteria and Fusobacteria (Fig. 1b).

At the genus level, *Bacteroides* was the dominant phylotype in both groups, but was significantly decreased in the liver cirrhosis group. Of the remaining genera, *Veillonella*, *Streptococcus*, *Clostridium* and *Prevotella* were enriched in the liver cirrhosis group, while *Eubacterium* and *Alistipes* were dominant in the healthy controls (Fig. 1a, b). The most abundant species in both liver cirrhosis and the healthy control groups were primarily from the *Bacteroides* genus. Of the 20 species that increased the most in abundance in the liver cirrhosis group, four were *Streptococcus* spp. and six were *Veillonella* spp., suggesting that the two genera might play an important role in liver cirrhosis. Of the species that decreased the most in abundance in the liver cirrhosis group, 12 were Bacteroidetes and seven were Firmicutes, specifically from the order Clostridiales.

Gut microbial species associated with cirrhosis

Our investigation included two phases. The first was discovery, where we compared 98 patients with liver cirrhosis and 83 healthy controls. The second was validation, with additional 25 patients and 31 controls. In the discovery phase, a Wilcoxon rank-sum test corrected for multiple testing by the Benjamini and Hochberg method was used to identify differentially abundant genes in patients and controls. At a stringent threshold (false discovery rate (FDR) < 0.0001), 75,245 genes were found: 49,830 were more abundant in the patients and 25,415 in the controls (Methods). Patients and controls could be clearly separated by principal component analysis based on the 75,245 genes; this was confirmed with the validation samples (Supplementary Table 3 and Extended Data Fig. 1c).

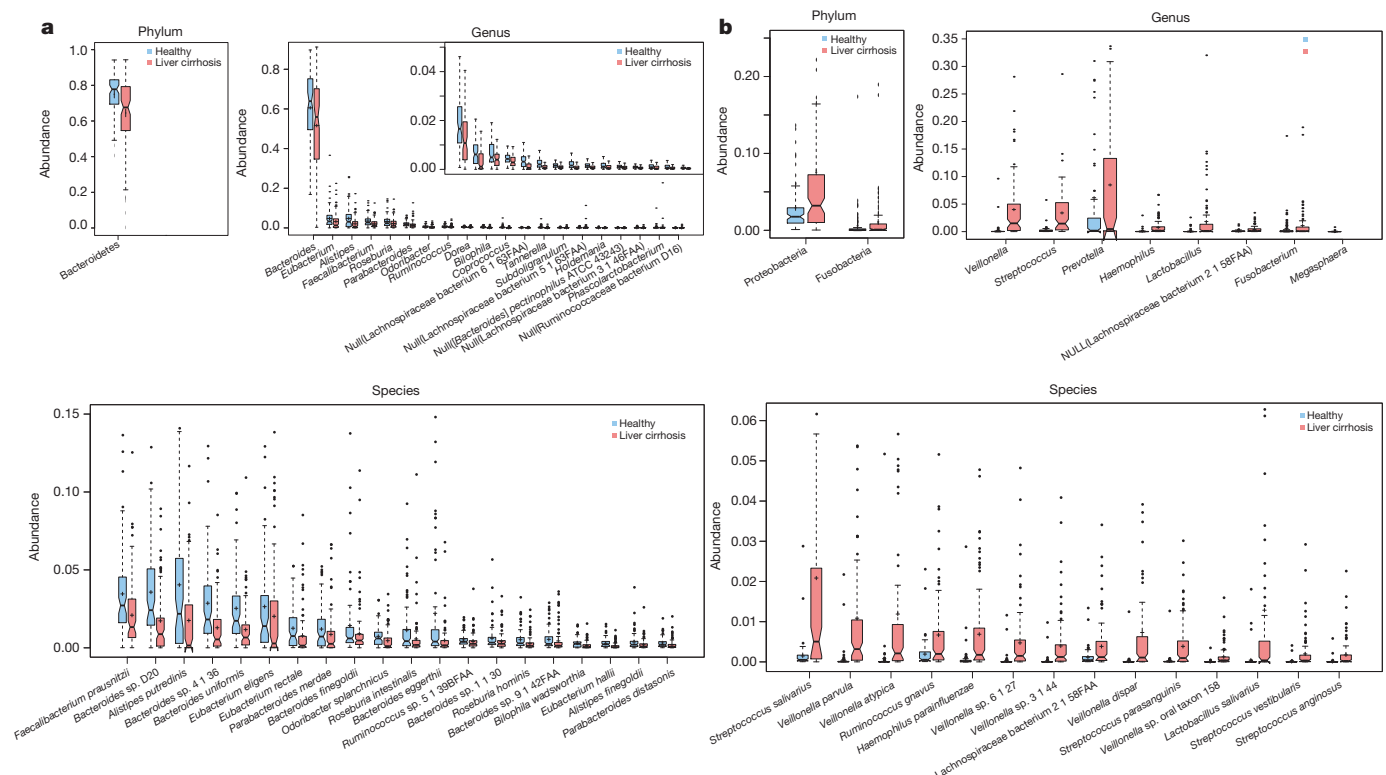


Figure 1 | Differentially abundant phyla in patients ($n = 98$) and healthy individuals ($n = 83$). The phylotypes decreased (a) and increased (b) in patients with liver cirrhosis at the phylum, genus and species levels. Blue and red represent healthy controls and patients with liver cirrhosis, respectively. Only the 20 most abundant species in each group are shown for clarity. The phylotypes with median relative abundances greater than 0.01% of total abundance in either the healthy control group or the liver cirrhosis group are

To explore further the microbial genes associated with liver cirrhosis we grouped them into clusters, denoted metagenomic species (MGS) here, on the basis of their abundance profiles^{27,30}. Of the 66 MGS, 38 and 28 were enriched in healthy individuals and patients, respectively. The significantly different abundance distribution between healthy and liver cirrhosis subjects is shown in Fig. 2 and Supplementary Table 4. A majority (82%) were also differentially abundant in the validation cohort ($q < 0.05$), in spite of the reduced statistical power due to the smaller cohort size.

Composition of bacterial communities varies considerably as a function of the overall gene richness^{27,28} and the loss of richness is associated with obesity and IBD^{27,28,31}. A large majority of the 38 MGS enriched in the healthy individuals (33, 86.8%) was correlated with the richness at $q < 10^{-3}$ in the Chinese cohort; 26 of these (78.8%) were similarly correlated in a Danish cohort (Extended Data Fig. 3). These observations indicate that gut communities of bacteria in healthy individuals across continents may be largely similar. Furthermore, gene richness was much lower in patients with liver cirrhosis than in healthy individuals (on average 389,000 and 497,000 genes, respectively; Supplementary Table 5 and Extended Data Fig. 4, top left). Interestingly, among the species enriched in healthy Chinese, were *Faecalibacterium prausnitzii*, which has anti-inflammatory properties and was found in a 'healthy' gene-rich microbiome^{27,28}, and *Coprococcus comes*, which might contribute to gut health through butyrate production. A similar butyrate production role may be played by three Lachnospiraceae and five Ruminococcaceae enriched in healthy individuals. A lower abundance of these species in patients with liver cirrhosis indicates that these individuals have a less healthy gut microbiome.

Most interestingly, a high proportion of MGS enriched in patients belong to taxa such as *Veillonella* ($n = 8$) or *Streptococcus* ($n = 6$), known to include species of oral origin (Supplementary Table 4). However, the small intestine also harbours such species³² and small-intestinal bacterial

included (FDR < 0.01, Wilcoxon rank-sum test corrected by the Benjamini and Hochberg method). The boxes represent the interquartile range (IQR), from the first and third quartiles, and the inside line represents the median. The whiskers denote the lowest and highest values within 1.5 IQR from the first and third quartiles. The circles represent outliers beyond the whiskers. The notches show the 95% confidence interval for the medians. If the notches of two boxes do not overlap, it gives evidence of a significant difference between the medians.

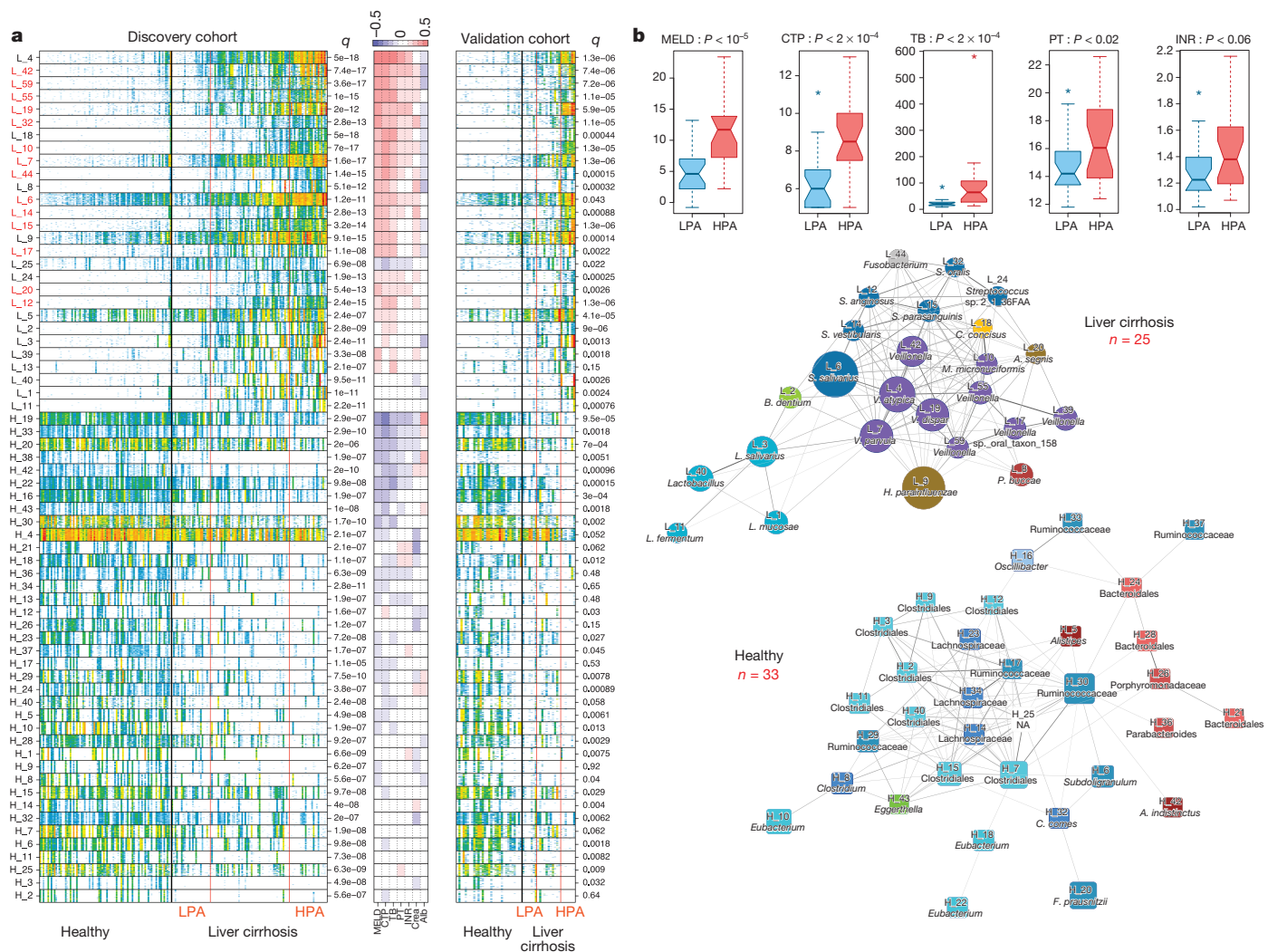


Figure 2 | Differentially abundant MGS in patients ($n = 123$) and healthy individuals ($n = 114$). **a**, Abundance of 50 'tracer' genes for each species in the discovery ($n_{\text{patients}} = 98$, $n_{\text{healthy}} = 83$) and validation cohorts ($n_{\text{patients}} = 25$, $n_{\text{healthy}} = 31$); oral species are highlighted in red. Genes are in rows, abundance is indicated by colour gradient (white, not detected; red, most abundant); the enrichment significance is shown (q indicates the Mann–Whitney P values corrected by the Benjamini and Hochberg method). Individuals are shown in columns, ordered by increasing abundance of patient-enriched species. Correlation of the species abundance and patients' clinical parameters in the discovery cohort are indicated in colour code (red and blue for positive and negative correlations; intensity reflects the level of correlation). MELD, model for end-stage liver disease; CTP, Child–Turcotte–Pugh score; TB, total bilirubin; PT, prothrombintime test; INR, international normalized ratio

overgrowth is frequently found in patients with liver cirrhosis³³. To explore the origin of the patient-enriched species, we used information from the HOMD³⁴ and GOLD³⁵ databases about the origin of the closely related sequenced isolates. We also constructed a catalogue of 114 publicly available genomes for *Streptococcus*, *Fusobacterium*, *Lactobacillus*, *Veillonella* and *Megasphaera* strains, originating mostly from mouth or gut (57 or 28, respectively; Supplementary Table 6) and used it for blastN and blastP analysis (Methods). Thirteen of the species were closest to an oral isolate whereas only six were closest to the gut isolates, a single species being from the ileum (Supplementary Table 4 and Extended Data Fig. 4, top right). Comparison with the three ileum metagenomes failed to reveal identity above that detected by comparison with the sequenced genomes (Methods). We conclude that oral commensals invade the gut in patients with liver cirrhosis. Possibly, an altered bile production in cirrhosis renders the gut more permissive and/or accessible to 'foreign'

describing coagulation of the blood in patients with liver cirrhosis; Crea, creatinine level; Alb, albumin level. **b**, Top, clinical parameters of patients for the lowest and highest patient-enriched species abundance (LPA and HPA, respectively; $n = 24$ for each). P values indicate the significance of the difference by Mann–Whitney U -test except MELD (Student's t -test). Middle and bottom, abundance-based species correlation network enriched in patients with liver cirrhosis ($n = 25$) and healthy individuals ($n = 33$), respectively. Two nodes are linked if the pooled variance z -test shows an FDR $< 10^{-9}$ when accounting for the compositionality effect (see Methods). The edge width is proportional to the correlation strength. The node size is proportional to the mean abundance in the respective population. Nodes with the same colour are classified in the same phylogenetic order level.

bacteria, as bile resistance may be required for survival in the human gut^{36,37}. As patient-enriched MGS include pathogens such as *Campylobacter* and *Haemophilus parainfluenzae*, these also might use the oral route to invade the gut, possibly via contaminated food. The invasion species foreign to the niche may occur not only in the colon but also in the ileum, and contribute to the small-intestinal bacterial overgrowth associated with liver cirrhosis. Among the patient-enriched species were *Streptococcus anginosus*, *Veillonella atypica*, *Veillonella dispar*, *Veillonella* sp. oral taxon and *Clostridium perfringens*, which have been reported to cause opportunistic infections^{38–40}.

To analyse the relations between the liver-cirrhosis-associated MGS, we generated networks based on co-abundance, for healthy individuals and patients with liver cirrhosis (Fig. 2b). A striking feature is that taxonomically related species tend to cluster, as reported previously²⁹. These observations indicate that the gut environment becomes permissive for

the development and maintenance of the related taxa in many individuals. Obviously, taxonomically unrelated species can also thrive in such environments, as observed with *Campylobacter concisus*, *H. parainfluenzae* or *Fusobacterium*, which tend to be associated with *Veillonella* in patients. The overall abundance of species enriched in patients reached high levels, exceeding 5% in over a quarter and approaching the extreme of 40%, whereas it was very low in healthy individuals (Extended Data Fig. 4, bottom). Interestingly, the severity of the disease was positively correlated with the abundance of a number of MGS enriched in patients and negatively correlated with those of the MGS enriched in controls (and therefore under-represented in patients; Fig. 2a). The disease status of the patients with the highest load of these bacteria was significantly worse than that of the patients with the lowest load (Fig. 2b, top). Such a 'dose response' is consistent with an active role of the enriched species in liver cirrhosis.

Microbial functions enriched in liver cirrhosis

To investigate the functional role of the gut microbiota in liver cirrhosis, we identified 4,801 KEGG (Kyoto Encyclopedia of Genes and Genomes database) orthologues and 13,970 eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups database) orthologues associated with the disease (Supplementary Tables 7 and 8). The most abundant KEGG orthologues in patients and controls were enzyme families. The most enriched orthologues in patients were membrane transport, similar to findings for IBDs^{19,20}, obesity⁴¹ and T2D²². In contrast, the most prevalent markers among the controls included those involved in carbohydrate metabolism, amino-acid metabolism, energy metabolism, signal transduction and the metabolism of cofactors and vitamins (Extended Data Fig. 5). At the module or pathway level, the liver-cirrhosis-associated markers included assimilation or dissimilation of nitrate to or from ammonia, denitrification, GABA (γ -aminobutyric acid) biosynthesis, GABA shunt, haem biosynthesis, phosphotransferase systems and some types of membrane transport, such as amino-acid transport. The control-enriched modules included histidine metabolism, ornithine biosynthesis, creatine pathway, carbohydrate metabolism, repair systems and glycosaminoglycan metabolism (Supplementary Table 9).

The enrichment of the modules for ammonia production in patients suggests a potential role of gut microbiota in hepatic encephalopathy, a complication related to liver cirrhosis that is characterized by hyperammonemia. Overproduction of ammonia by gut bacteria might contribute to increased levels of ammonia in blood. Manganese-related transport system modules enriched in patients possibly contribute to the changes in concentrations of manganese. The accumulation of manganese within the basal ganglia in patients with end-stage liver disease may have a role in the pathogenesis of chronic hepatic encephalopathy⁴², a main complication of liver cirrhosis. The hydrodynamic venous shunt and liver failure could promote this accumulation, which, in turn, causes metabolic disorders of the nerve cell enzymes, affects transmission function of neural synapses and eventually leads to hepatic encephalopathy⁴⁰. Finally, the modules for GABA biosynthesis were enriched in the patients. The GABA neurotransmitter system is involved in the pathogenesis of hepatic encephalopathy in humans⁴³. Because of the hydrodynamic venous shunt and liver failure, GABA levels in the blood are increased⁴⁴, and could go through the blood-brain barrier to activate GABA receptor and cause hepatic encephalopathy. Microbiome modulation, aiming at manganese elimination and lowering of GABA levels in the gut, might provide a new therapeutic option for the treatment of hepatic encephalopathy.

Microbial dysbiosis in chronic diseases

It is unclear whether a gut microbial dysbiosis in type 2 diabetes (T2D)²², IBD⁴¹ and liver cirrhosis¹³ is similar or unique for each disease. We compared the differences between the gut microbiota from patients with liver cirrhosis, T2D and IBD, and organized the disease-associated gene, KEGG orthologue group and eggNOG orthologue group markers into patient- and control-enriched groups. We then identified markers common

to different disease pairs (T2D and liver cirrhosis, liver cirrhosis and IBD, and IBD and T2D) and to the three diseases (Supplementary Table 10). Different diseases displayed a relatively unique profile, even if some markers were shared (Extended Data Fig. 6a, b). Most liver-cirrhosis-enriched markers had low *P* values (Extended Data Fig. 6c), implying that patients with liver cirrhosis had more severe dysbiosis than patients with T2D. Functional differences between liver cirrhosis and T2D were also detected at the pathway level, even if there was a significant increase in membrane transport markers in both (Extended Data Figs 7 and 8). Most functional markers in both diseases were from categories of carbohydrate metabolism, metabolism of cofactors and vitamins, amino-acid metabolism and signal transduction. In contrast, most cell motility markers in the KEGG orthologue group were enriched in liver cirrhosis or T2D but not both, possibly indicating a unique role in each disease (Extended Data Fig. 8a, b). However, similar cell motility markers and pathways in the KEGG orthologue group were enriched both in liver cirrhosis and in T2D controls, suggesting a possible role in health (Extended Data Figs 8c, d and 9a, b).

Gene markers that identify patients with liver cirrhosis

We used a pattern recognition technique to identify patients by gut microbiota information in the discovery cohort ($n = 181$). For this we selected 46,000 genes, half enriched in patients and half in controls (Supplementary Table 11). From this set we selected 15 optimal gene markers by a minimum redundancy-maximum relevance (mRMR) method combined with an incremental feature search, which showed the highest value of Matthews correlation coefficient (Extended Data Fig. 9c). A support vector machine discriminator was constructed using the same samples and 15 gene markers (Supplementary Table 12), with the training and leave-one-out cross-validation AUC (area under the receiver operating characteristic curve) achieving 0.918 (confidence interval: 0.881–0.955) (Fig. 3b) and 0.838, respectively. The validation cohort of 31 healthy controls and 25 patients with liver cirrhosis showed an AUC value of 0.836 (95% confidence interval 0.730–0.943) (Fig. 3c) for these samples, confirming that the gut microbiota information could be applied to identify patients accurately.

To facilitate the clinical application of the 15 optimal gene markers, we propose a patient discrimination index (PDI). The high correlation coefficient value between the ratio of patients in our cohort and the PDI (Fig. 3a and Supplementary Table 13) indicates that the PDI could be used to identify patients with liver cirrhosis. The discriminatory power of the PDI was then validated using an independent group (Fig. 3d). The average PDI index between the control and the patient groups was significantly different ($P < 8.18 \times 10^{-5}$, Wilcoxon rank-sum test), confirming the potential use of gut microbiota information for identifying patients with liver cirrhosis.

Discussion

To study gut microbiota in liver cirrhosis we first established a novel gut gene catalogue (liver cirrhosis catalogue), including 98 patients with liver cirrhosis and 83 healthy control individuals. Comparison with the previously established MetaHIT and T2D²² gene catalogues indicated a common core of approximately 800,000 genes and a considerable proportion of catalogue-specific genes (37.01% of MetaHIT, 36.59% of T2D and 18.02% of liver cirrhosis), indicating that the current gene sets are still limited and should be completed by inclusion of more individuals. Interestingly, although the T2D and liver cirrhosis gene sets are both derived from Chinese populations, the number of unique genes in each gene set was large. This might be due to the difference in disease profiles and to the different genotypes, body mass indices, age⁴⁵ and dietary habits⁴⁶ (Supplementary Table 14 and Extended Data Fig. 10). Nevertheless, there was no significant difference in the abundance of main phyla ($P > 0.01$); of the top 30 most abundant genera and species, 28 and 26, respectively, were the same in both studies, and there were no significant differences in abundance for most of them. Furthermore, the top four species were exactly the same. These results, and the similarity of

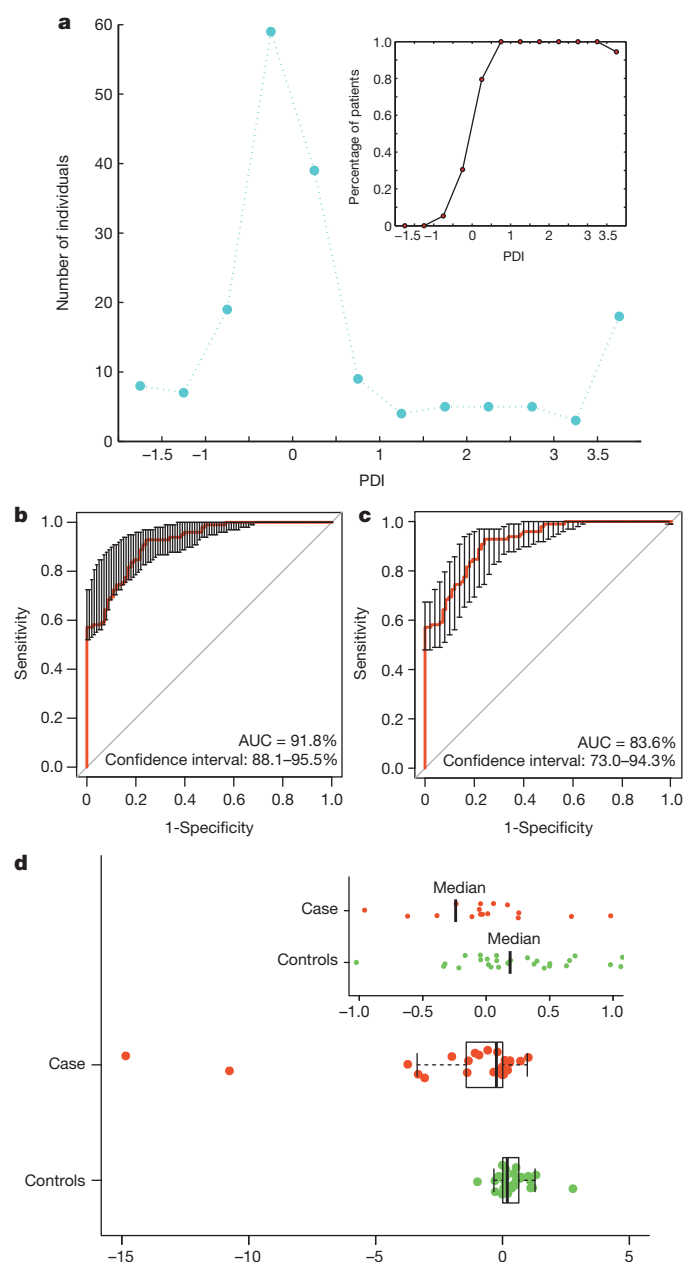


Figure 3 | PDI on the basis of gut microbial biomarkers. **a**, A PDI was calculated for each individual from 15 gene markers selected using the mRMR approach to evaluate the risk of liver cirrhosis. The filled blue circles show the distribution of liver cirrhosis indices for all individuals (bins of 0.5 PDI units were used; values less than -1.5 and greater than 3.5 were grouped). Inset, the proportion of patients with liver cirrhosis in the corresponding bins. **b**, **c**, The AUC is shown for the training (**b**) and validation (**c**) samples. **d**, The liver cirrhosis PDI was computed for an additional 25 liver cirrhosis samples and 31 healthy control samples. The box depicts the interquartile range between the first and third quartiles (25th and 75th percentiles, respectively); the line inside denotes the median. Inset, the PDI without the outliers.

controls with the healthy Danish population, point towards overall similarity of the microbiota in healthy individuals.

Use of the liver cirrhosis gene catalogue, in conjunction with the quantitative metagenomics approach, revealed a major change of the gut microbiota in the patients with liver cirrhosis, mainly because of a massive invasion of the gut by oral bacterial species. Correlation of the severity of the disease with the abundance of the invading species suggests that they may play an active role in the pathology. This was not noted in a previous study, where the 16S-based approach probably lacked the required

species-level resolution, even if similar trends in taxonomy change between the liver cirrhosis group and the healthy controls at the phylum, class and order levels were observed¹³. Some of the MGS depleted in patients were negatively associated with the severity of the disease (Fig. 2). This opens avenues to the development of novel probiotics, which might help combat the aggravation of liver cirrhosis. More generally, modulation of microbiota to correct the major dysbioses we report might open new avenues to treatment of liver cirrhosis.

A combination of 15 microbial genes discriminates patients with liver cirrhosis from healthy individuals, with a high specificity. This could lead to a new way of monitoring and preventing liver cirrhosis. None of the 15 markers found in the liver cirrhosis study overlapped with the 50 markers found in the T2D study²², indicating that diagnosis of different diseases with microbiota-targeted biomarkers may be a powerful tool for disease detection.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 April 2013; accepted 9 June 2014.

Published online 23 July 2014.

- Fouts, D. E., Torralba, M., Nelson, K. E., Brenner, D. A. & Schnabl, B. Bacterial translocation and changes in the intestinal microbiome in mouse models of liver disease. *J. Hepatol.* **56**, 1283–1292 (2012).
- Cesaro, C. *et al.* Gut microbiota and probiotics in chronic liver diseases. *Digest. Liver Dis.* **43**, 431–438 (2011).
- Wiest, R. & Garcia-Tsao, G. Bacterial translocation (BT) in cirrhosis. *Hepatology* **41**, 422–433 (2005).
- Nolan, J. P. The role of intestinal endotoxin in liver injury: a long and evolving history. *Hepatology* **52**, 1829–1835 (2010).
- Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312**, 1355–1359 (2006).
- Garcia-Tsao, G. & Wiest, R. Gut microflora in the pathogenesis of the complications of cirrhosis. *Best Pract. Res. Clin. Gastroenterol.* **18**, 353–372 (2004).
- Wiest, R., Krag, A. & Gerbes, A. Spontaneous bacterial peritonitis: recent guidelines and beyond. *Gut* **61**, 297–310 (2012).
- Bass, N. M. *et al.* Rifaximin treatment in hepatic encephalopathy. *N. Engl. J. Med.* **362**, 1071–1081 (2010).
- Benten, D. & Wiest, R. Gut microbiome and intestinal barrier failure—the “Achilles heel” in hepatology? *J. Hepatol.* **56**, 1221–1223 (2012).
- Yan, A. W. *et al.* Enteric dysbiosis associated with a mouse model of alcoholic liver disease. *Hepatology* **53**, 96–105 (2011).
- De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl Acad. Sci. USA* **107**, 14691–14696 (2010).
- Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nature Rev. Genet.* **13**, 260–270 (2012).
- Chen, Y. *et al.* Characterization of fecal microbial communities in patients with liver cirrhosis. *Hepatology* **54**, 562–572 (2011).
- Nelson, K. E. *et al.* A catalog of reference genomes from the human microbiome. *Science* **328**, 994–999 (2010).
- Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023 (2006).
- Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
- Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
- Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proc. Natl Acad. Sci. USA* **102**, 11070–11075 (2005).
- Lepage, P. *et al.* Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* **141**, 227–236 (2011).
- Garrett, W. S. *et al.* Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8**, 292–300 (2010).
- Wen, L. *et al.* Innate immunity and intestinal microbiota in the development of type 1 diabetes. *Nature* **455**, 1109–1113 (2008).
- Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
- Vijay-Kumar, M. *et al.* Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* **328**, 228–231 (2010).
- Karlsson, F. H. *et al.* Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nature Commun.* **3**, 1245 (2012).
- The Human Microbiome Project Consortium. A framework for human microbiome research. *Nature* **486**, 215–221 (2012).
- The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
- Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).

28. Cotillard, A. *et al.* Dietary intervention impact on gut microbial gene richness. *Nature* **500**, 585–588 (2013).
29. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
30. Nielsen, H. B. *et al.* Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnol.* <http://dx.doi.org/10.1038/nbt.2939> (2014).
31. Albersen, M. *et al.* Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotechnol.* **31**, 533–538 (2013).
32. Zoetendal, E. G. *et al.* The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J.* **6**, 1415–1426 (2012).
33. Bauer, T. M. *et al.* Small intestinal bacterial overgrowth in human cirrhosis is associated with systemic endotoxemia. *Am. J. Gastroenterol.* **97**, 2364–2370 (2002).
34. Chen, T. *et al.* The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* **2010**, baq013 (2010).
35. Pagani, I. *et al.* The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* **40**, D571–D579 (2012).
36. Saarela, M., Mogensen, G., Fonden, R., Matto, J. & Mattila-Sandholm, T. Probiotic bacteria: safety, functional and technological properties. *J. Biotechnol.* **84**, 197–215 (2000).
37. Merritt, M. E. & Donaldson, J. R. Effect of bile salts on the DNA and membrane integrity of enteric bacteria. *J. Med. Microbiol.* **58**, 1533–1541 (2009).
38. Marchandin, H. *et al.* Prosthetic joint infection due to *Veillonella dispar*. *Eur. J. Clin. Microbiol. Infect. Dis.* **20**, 340–342 (2001).
39. Hwang, J. J., Lau, Y. J., Hu, B. S., Shi, Z. Y. & Lin, Y. H. *Haemophilus parainfluenzae* and *Fusobacterium necrophorum* liver abscess: a case report. *J. Microbiol. Immunol. Infect.* **35**, 65–67 (2002).
40. Xu, M. *et al.* Changes of fecal *Bifidobacterium* species in adult patients with hepatitis B virus-induced chronic liver disease. *Microb. Ecol.* **63**, 304–313 (2012).
41. Greenblum, S., Turnbaugh, P. J. & Borenstein, E. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc. Natl Acad. Sci. USA* **109**, 594–599 (2012).
42. Krieger, D. *et al.* Manganese and chronic hepatic encephalopathy. *Lancet* **346**, 270–274 (1995).
43. Ferenci, P., Schafer, D. F., Kleinberger, G., Hoofnagle, J. H. & Jones, E. A. Serum levels of gamma-aminobutyric-acid-like activity in acute and chronic hepatocellular disease. *Lancet* **ii**, 811–814 (1983).
44. Minuk, G. Y., Winder, A., Burgess, E. D. & Sarjeant, E. J. Serum gamma-aminobutyric acid (GABA) levels in patients with hepatic encephalopathy. *Hepatology* **32**, 171–174 (1985).
45. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
46. Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by the National Program on Key Basic Research Project (2013CB531401), the National Natural Science Foundation of China (81301475 and 81330011), the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (81121002), the Technology Group Project for Infectious Disease Control of Zhejiang Province (2009R50041) and the Metagenopolis grant ANR-11-DPBS-0001. We thank Q. Cao, K. Su, J. Shao and A. Ghazlane for help with data computation, and H. Zhang, H. Lu, Q. Bao, J. Ge, J. Jiang, Z. Ren and M. Ye for assistance with sample collection. We are thankful to the MetaHIT consortium for generating the gut gene set and the Human Microbiome Project for generating the reference genomes from human gut microbes.

Author Contributions L.J.L., S.D.E., S.S.Z. and N.Q. designed the project. L.J.L., S.P.K. and N.Q. managed the project. F.L.Y., N.Q., Y.F.C., J.G., G.R.Q., X.J.H. and B.W.Z. collected samples and performed clinical study. J.G., Y.T.C. and W.X. performed DNA extraction experiments. Y.J., L.J.W., J.W.Z. and S.J.N. performed library construction and sequencing. L.J.L. and S.D.E. designed the analysis. N.Q., A.L., E.P., E.L.C., L.L., N.P., P.L., J.M.B., C.H.Y. and W.C.D. analysed the data. A.L. and N.Q. did the functional annotation analyses. L.S., E.P., E.L.C. and A.L. analysed the statistics. N.Q., F.L.Y., L.S. and E.P. wrote the paper. L.J.L. and S.D.E. revised the paper.

Author Information The raw Illumina read data for all samples have been deposited in the European Bioinformatics Institute European Nucleotide Archive under accession number ERP005860. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.J.L. (ljli@zju.edu.cn), S.S.Z. (zyzsss@zju.edu.cn) or S.D.E. (dusko.ehrlich@jouy.inra.fr).

RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer

Andrew L. Wolfe^{1,2*}, Kamini Singh^{1*}, Yi Zhong³, Philipp Drewe³, Vinagolu K. Rajasekhar⁴, Viraj R. Sanghvi¹, Konstantinos J. Mavrikis^{1†}, Man Jiang¹, Justine E. Roderick⁵, Joni Van der Meulen^{1,6}, Jonathan H. Schatz^{1,7†}, Christina M. Rodrigo⁸, Chunying Zhao¹, Pieter Rondou⁶, Elisa de Stanchina⁹, Julie Teruya-Feldstein¹⁰, Michelle A. Kelliher⁵, Frank Speleman⁶, John A. Porco Jr⁸, Jerry Pelletier^{11,12,13}, Gunnar Räscher³ & Hans-Guido Wendel¹

The translational control of oncoprotein expression is implicated in many cancers. Here we report an eIF4A RNA helicase-dependent mechanism of translational control that contributes to oncogenesis and underlies the anticancer effects of silvestrol and related compounds. For example, eIF4A promotes T-cell acute lymphoblastic leukaemia development *in vivo* and is required for leukaemia maintenance. Accordingly, inhibition of eIF4A with silvestrol has powerful therapeutic effects against murine and human leukaemic cells *in vitro* and *in vivo*. We use transcriptome-scale ribosome footprinting to identify the hallmarks of eIF4A-dependent transcripts. These include 5' untranslated region (UTR) sequences such as the 12-nucleotide guanine quartet (CGG)₄ motif that can form RNA G-quadruplex structures. Notably, among the most eIF4A-dependent and silvestrol-sensitive transcripts are a number of oncogenes, superenhancer-associated transcription factors, and epigenetic regulators. Hence, the 5' UTRs of select cancer genes harbour a targetable requirement for the eIF4A RNA helicase.

The activation of translation contributes to malignant transformation and is an emerging target for cancer therapies^{1,2}. For example, oncogenic RAS, ERK and AKT stimulate cap-dependent translation via mTORC1 and the eIF4E-binding protein (4E-BP)³. The initiation of cap-dependent translation involves a tightly controlled multi-protein initiation complex⁴. This includes the cap-binding protein eIF4E, which is abundantly expressed in cancer and can transform fibroblasts and promote tumour development *in vivo*^{1,5–8}. Another key factor is the eIF4A RNA helicase, which has a role in scanning the messenger RNA 5' UTR for a translation start site and has been implicated in the translation of mRNAs with long and complex 5' UTRs⁹. EIF4A is the molecular target for three distinct natural compounds—silvestrol, hippuristanol and pateamine A—and these compounds show promising anticancer activity^{10,11}. Recent studies have defined the helicase action of eIF4A, as well as its protein interactions¹², and revealed eIF4A to have a role in microRNA action¹³. Exactly how eIF4A affects translation and which specific mRNA features necessitate its helicase activity have not been defined.

Ribosome footprinting provides a snapshot of translation across the transcriptome¹⁴. It uses deep sequencing to identify ribosome-protected RNA fragments and compare them to total transcript levels. Among the various applications of this technology, the most relevant to this study concern the translational effects of mTORC1 inhibition^{15,16}. We use ribosome footprinting to identify eIF4A-dependent transcripts and their hallmark features.

eIF4A is an oncogene in T-ALL

NOTCH-driven T-cell acute lymphoblastic leukaemia (T-ALL) exemplifies the activation of AKT/mTORC1 and cap-dependent translation

in cancer. In 36 paediatric T-ALL samples we find *PTEN* mutations (14%) and deletions (11%)¹⁷, *NOTCH1* mutations (56%)¹⁸ and an *IL7R* mutation (3%)¹⁹ (Extended Data Fig. 1a–c and Supplementary Table 1). These genetic lesions promote T-ALL development in a murine T-ALL model (Fig. 1a)²⁰. *Notch* causes T-ALL with a mean latency of 91.5 days ($n = 14$), knockdown of *Pten* or expression of the mutant *Il7r/p.Leu242_Leu243insAsnProCys* accelerates disease onset (*Pten*, 47.1 days, $n = 10$, $P < 0.0001$; *Il7r*, 35.5 days, $n = 4$, $P < 0.0001$). Remarkably, expression of eIF4E or eIF4A1 similarly accelerates leukaemia development (eIF4E, 30.75 days, $n = 4$, $P < 0.0001$; eIF4A1, 33.8 days, $n = 5$, $P < 0.0001$) (Fig. 1b and Extended Data Fig. 1d). All T-ALLs are CD4/CD8 double positive, and increased ribosomal S6 phosphorylation indicates mTORC1 activation in *Pten*-deficient and *Il7r*-expressing T-ALLs (Extended Data Fig. 1d, f–i). EIF4E and eIF4A1 are required to maintain T-ALL and cells expressing a constitutive 4E-BP1-encoding allele (4E-BP1(4A))²¹ or an eIF4A1 knockdown construct are rapidly eliminated from mixed populations (Fig. 1c, d and Extended Data Fig. 1e) ($P = 0.000002$, vector versus 4E-BP1(4A); $P = 0.000008$, vector versus short hairpin RNA (shRNA)-targeted eIF4A (shEIF4A)).

Silvestrol and a synthetic analogue, (±)-CR-31-B, inhibit eIF4A1/2 (Extended Data Fig. 2a, b)^{11,22}. A reporter assay confirms that both drugs preferentially block cap-dependent translation of *Renilla* luciferase compared to firefly luciferase expressed from the hepatitis C virus (HCV) internal ribosome entry site (IRES) (Fig. 2a and Extended Data Fig. 2c). Silvestrol induces cell death in primary human T-ALL samples, cell lines, and murine T-ALLs at nanomolar half-maximum inhibitory concentration (IC₅₀) values (Fig. 2b and Extended Data Fig. 2d, e). *In vivo*, silvestrol is effective against murine or xenografted T-ALLs

¹Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ²Weill Cornell Graduate School of Medical Sciences, New York, New York 10065, USA.

³Computational Biology Department, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ⁴Stem Cell Center and Developmental Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ⁵Department of Cancer Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01605 USA. ⁶Center for Medical Genetics, Ghent University Hospital, De Pintelaan 185, B-9000 Ghent, Belgium. ⁷Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ⁸Department of Chemistry, Center for Chemical Methodology and Library Development, Boston University, Boston, Massachusetts 02215, USA. ⁹Molecular Pharmacology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ¹⁰Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ¹¹Department of Biochemistry, McGill University, Montreal, Quebec H3G 1Y6, Canada. ¹²Department of Oncology, McGill University, Montreal, Quebec H3G 1Y6, Canada. ¹³The Rosalind and Morris Goodman Cancer Research Center, McGill University, Montreal, Quebec H3G 1Y6, Canada.

[†]Present addresses: Novartis, Cambridge, Massachusetts 02139, USA (K.J.M.); The University of Arizona Cancer Center, Tucson, Arizona 85719, USA (J.H.S.).

*These authors contributed equally to this work.

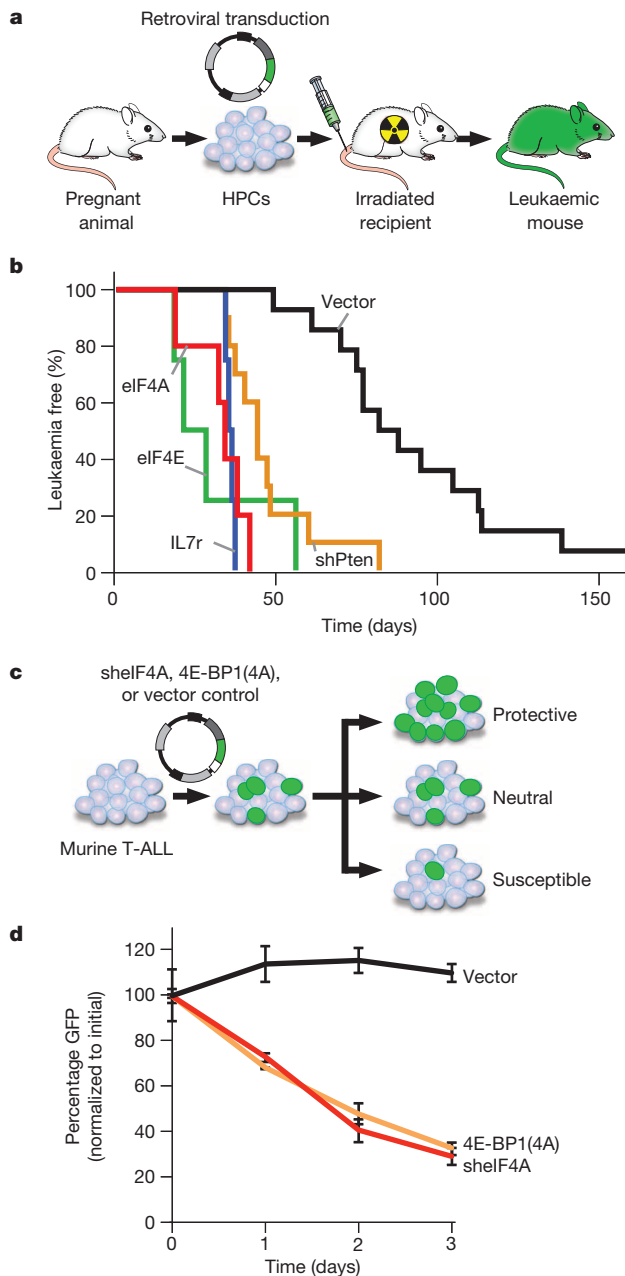


Figure 1 | eIF4A promotes T-ALL development in vivo. **a**, Diagram of the intracellular NOTCH1 (ICN)-driven murine T-ALL model. HPCs, hematopoietic progenitor cells. **b**, Kaplan–Meier analysis showing time to leukaemia development after transplantation of HPCs transduced with NOTCH1-ICN and empty vector (black, $n = 14$), eIF4E (green, $n = 4$), eIF4A1 (red, $n = 5$), IL7r p.Leu242_Leu243insAsnProCys (blue, $n = 4$), shPten (orange, $n = 10$). **c**, Experimental design of competition experiments. **d**, Results as percentage of each starting GFP-positive population of murine T-ALL cells partially transduced with vector/GFP, shPten or 4E-BP1(4A). CR, (\pm) -CR-31-B. Mean and standard deviation (s.d.) are shown, $n = 3$ biological replicates.

(Fig. 2c and Extended Data Fig. 2f–h). In KOPT-K1 tumour-bearing ($\sim 1 \text{ cm}^3$) NOD/SCID mice, treatment with silvestrol (0.5 mg kg^{-1} , intraperitoneal (i.p.), days 0–6, $n = 7$, $P < 0.001$) or (\pm) -CR-31-B (0.2 mg kg^{-1} , i.p., days 0–6, $n = 8$, $P < 0.001$) delays tumour growth, and causes apoptosis and cell cycle arrest (Fig. 2c, d and Extended Data Fig. 2g, h). Detailed toxicology shows that this treatment is well-tolerated in mice (Extended Data Fig. 3a–j and Supplementary Table 2). Rapamycin induces an S6 kinase-dependent feedback activation of AKT (T308)²³; by contrast, silvestrol or (\pm) -CR-31-B does not trigger this

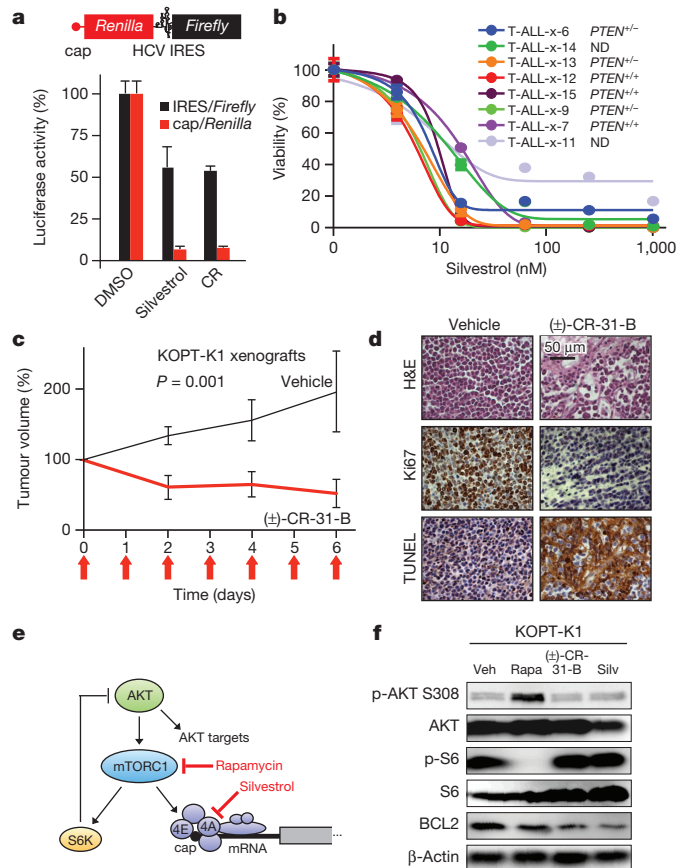


Figure 2 | Silvestrol has single-agent activity against T-ALL. **a**, Reporter system with capped *Renilla* luciferase (red) and *Firefly* luciferase under the HCV IRES (black). Bottom, relative levels of *Renilla* luciferase (red) and *Firefly* (black) luciferase upon vehicle (dimethylsulphoxide (DMSO)), silvestrol, or (\pm) -CR-31-B treatment. Mean and s.d. is shown; $n = 3$ biological replicates. **b**, Viability of primary patient T-ALL samples treated with silvestrol (48 h; mean and s.d. of 4 replicates is shown). ND, not done. **c**, Tumour size of KOPT-K1 xenografts treated with (\pm) -CR-31-B (0.2 mg kg^{-1}) or vehicle. Mean and s.d. of 5 tumours is shown. **d**, Immunohistochemical analysis of (\pm) -CR-31-B-treated KOPT-K1 tumours. H&E, haematoxylin and eosin; TUNEL, TdT-mediated dUTP nick end labelling. **e**, Diagram of drug targets. **f**, Lysates of KOPT-K1 cells treated with vehicle (Veh), rapamycin (Rapa; 25 nM), (\pm) -CR-31-B (CR; 25 nM), or silvestrol (Silv; 25 nM) for 48 h and probed as indicated. p, phosphorylated.

response in KOPT-K1 cells (Fig. 2e, f). This result implies that inhibition of eIF4A is effective without having an effect on S6 kinase.

Ribosome footprinting

For footprinting studies, we treated KOPT-K1 cells with 25 nM of silvestrol or vehicle for 45 min, then deep-sequenced total RNA and ribosome-protected RNA (Fig. 3a)¹⁴. We removed reads mapping to ribosomal RNAs, non-coding RNAs, library linkers and incomplete alignments (Extended Data Fig. 4a, b). The majority of the remaining reads between 25–35 nucleotides in length mapped to protein-coding genes (Extended Data Fig. 4c, d). The total number of ribosome footprint reads that mapped to exons was 3.2 million in control-treated and 3.4 million in silvestrol-treated samples, and this corresponded to $\sim 11,128$ protein-coding genes. Ribosome footprint reads showed a wider variation between control and silvestrol groups than the total RNA sequences, indicating minimal transcriptional variation (Extended Data Fig. 4e). The number of ribosomes occupying any transcript is given as gene-specific ribosome footprint reads per one million total reads (RPM). The RPM frequency distribution in control and silvestrol samples was overlapping, indicating that silvestrol equally affected mRNAs with high

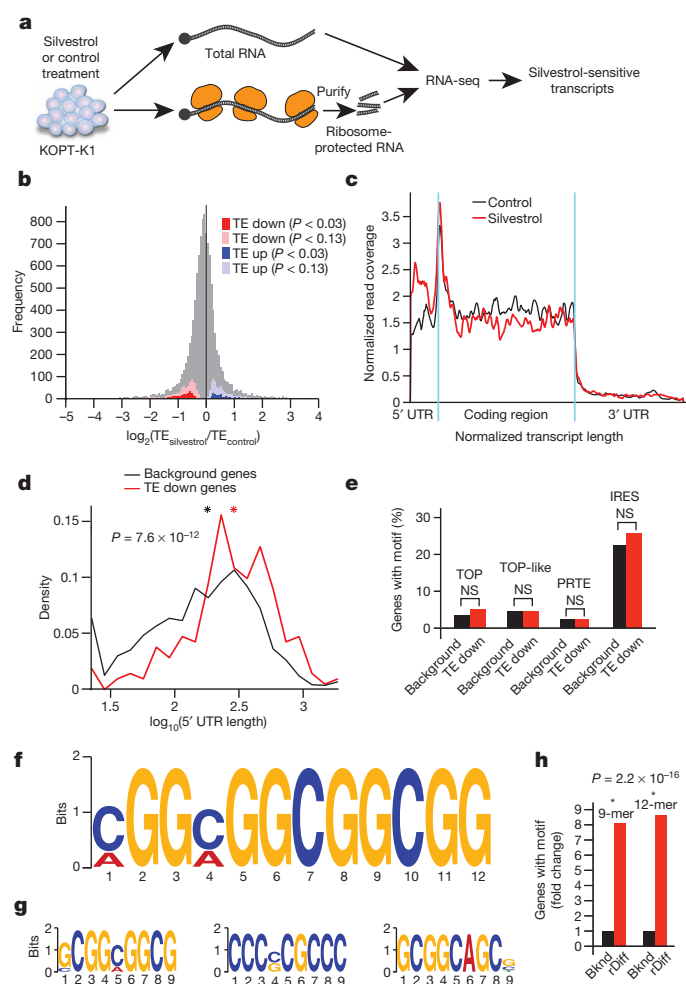


Figure 3 | Ribosome footprinting defines the effects of silvestrol on translation. **a**, Schematic of the ribosome footprinting study. RNA-seq, RNA-sequencing. **b**, Frequency distribution of the ratio of TE in control and silvestrol-treated samples ($TE_{\text{silvestrol}}/TE_{\text{control}}$). More or less affected mRNAs are identified as TE down (red) and TE up (blue), respectively; $n = 2$ replicates. **c**, Ribosome distribution for 62 TE down and rDiff-positive transcripts upon silvestrol (red) or vehicle (black) treatment. Ribosome footprint coverage and transcript length are normalized, blue indicates translation start and stop sites. **d**, Comparison of 5' UTR lengths for TE down versus background genes. Mathematical density is scaled such that all values on the x-axis sum to 1. Red, TE down; black, background genes. Asterisks indicate mean value, $n = 2$ replicates. **e**, Prevalence of the indicated 5' UTR motifs among the TE down and background genes. NS, not significant ($P > 0.05$). **f**, Twelve-nucleotide motif enriched in TE down genes ($P = 2 \times 10^{-16}$). **g**, The three most common 9-nucleotide motifs in TE down genes. **h**, Enrichment of 12-nucleotide (12-mer) and 9-nucleotide (9-mer) motifs in the rDiff gene set. Bknd, background.

and low ribosome occupancy (Extended Data Fig. 4f). Polysome analysis and metabolic labelling with L-azidohomoalanine (AHA) confirmed an inhibitory effect on translation (AHA: silvestrol $\sim 60\%$, $P = 3.6 \times 10^{-3}$, silvestrol versus vehicle; cycloheximide 80% , $P = 2 \times 10^{-4}$, cycloheximide versus vehicle) (Extended Data Fig. 4g, h). The translational efficiency (TE) for each mRNA is calculated by normalizing the ribosome footprint frequency to transcript length and total transcript abundance (reads per kilobase per million reads (RPKM)). RPKM values for ribosome footprints from vehicle and silvestrol samples were correlated ($R^2 = 0.94$), indicating an overall inhibitory effect (Extended Data Fig. 4i).

We developed the differential-expression-normalized ribosome occupancy (DERseq; based on DEXseq²⁴) algorithm to identify mRNAs that were most strongly affected by silvestrol. We used a cut-off at $P < 0.03$ (Z -score > 2.5) to define groups of mRNAs whose translational efficiency

was the most (TE down; red) or least (TE up; blue) affected by silvestrol compared to background (grey) (Fig. 3b and Supplementary Table 3a–c). The TE down group includes 281 mRNAs (220 with annotated 5' UTRs), TE up includes 190 mRNAs, and the background list includes 2,243 mRNAs. The footprinting methodology also provides positional information and the rDiff algorithm identifies mRNAs with significant shifts in ribosome footprint distribution²⁵. For example, silvestrol caused an accumulation of ribosome footprints in the 5' UTR of 847 protein-coding transcripts (rDiff positive genes; 641 with annotated 5' UTRs, $P < 0.001$) (Supplementary Table 3d). Sixty-two transcripts showed both decreased TE and altered ribosome footprint distribution, whereas we observed no change in distribution among TE up genes (Fig. 3c, Extended Data Fig. 5a–c and Supplementary Table 3e).

The (CGG)₄ motif marks eIF4A-dependent mRNAs

We compared the TE up, TE down and background groups (Supplementary Table 3a–c), and confirm that mRNAs with longer 5' UTRs were significantly enriched in the TE down group (mean UTR length, 368 nucleotides in TE down; 250 nucleotides in background, $P = 7.6 \times 10^{-12}$, silvestrol versus control, by two-sample Kolmogorov–Smirnov) (Fig. 3d). The TE up group showed no difference in 5' UTR length (265 nucleotides; $P = 0.165$, silvestrol versus control) (Extended Data Fig. 4j). We found no predilection for known regulatory elements in the TE down group; these included TOP²⁶, TOP-like sequences¹⁵, IRES²⁷ and pyrimidine-rich translational elements (PRTes)²⁶ (Fig. 3e). The TE up group showed the expected enrichment in IRES elements (Extended Data Fig. 4k). Using the DREME algorithm²⁸, we found a striking enrichment of a 12-nucleotide (CGG)₄ motif and related but shorter 9-nucleotide motifs among the 220 TE down genes (Fig. 3f, g and Supplementary Table 4a, b). Both the 12-nucleotide and 9-nucleotide motifs were significantly enriched over background and their frequency was significantly higher than expected merely by the longer UTR length ($P < 2.2 \times 10^{-16}$, one-sided binomial test) (Fig. 3h). We found no enriched motif in the TE up group. Transcripts with an altered ribosome footprint distribution (rDiff positive) showed similar features. These included longer 5' UTRs (rDiff, mean 271 nucleotides versus background: mean 230 nucleotides; $P = 0.004$) (Extended Data Fig. 5d) and absence of enrichment for TOP, TOP-like, PRTes or IRES elements (Extended Data Fig. 5e). Notably, we again identified the 12-nucleotide motif seen in the TE analysis but shifted by one nucleotide (Extended Data Fig. 5f and Supplementary Table 4c), along with 9-nucleotide variants, highly significantly enriched in the rDiff positive set compared to background ($P = 2.2 \times 10^{-16}$) (Extended Data Fig. 5g and Supplementary Table 4d).

The (CGG)₄ motif corresponds to G-quadruplexes

We noticed that in many instances the 5' UTR motifs coincided with computationally predicted G-quadruplex structures²⁹. For example, 51% of the 12-nucleotide (CGG)₄ sequences and 43% of the most common 9-nucleotide motifs localized precisely to the G-quadruplex structures—the *ADAM10* 5' UTR provides an example (Fig. 4a, b, Extended Data Fig. 6a and Supplementary Table 4e–k). G-quadruplex structures form by non-Watson–Crick interactions between paired guanine quartets connected by at least one linker nucleotide (A or C) in a parallel or anti-parallel orientation. Accordingly, G-quadruplex structures were significantly enriched among TE down genes and 79 of 220 TE down transcripts harboured at least one G-quadruplex ($P = 2 \times 10^{-11}$) (Fig. 4a–c and Supplementary Table 4e–k).

We confirmed by circular dichroism that the 12-nucleotide and extended 9-nucleotide sequence motifs form G-quadruplexes. Briefly, we compared the molar ellipticity of RNA oligomers encoding the 12-nucleotide and the 9-nucleotide motifs to a known G-quadruplex element in human telomeric RNA³⁰ and a randomly organized oligomer with equal GC content and length. The 9-nucleotide sequences included either two (oligonucleotide 1) or five (oligonucleotide 2) flanking nucleotides as seen in the 5' UTRs of *MTA2*, *TGFB1*, *MAPKAP1* and

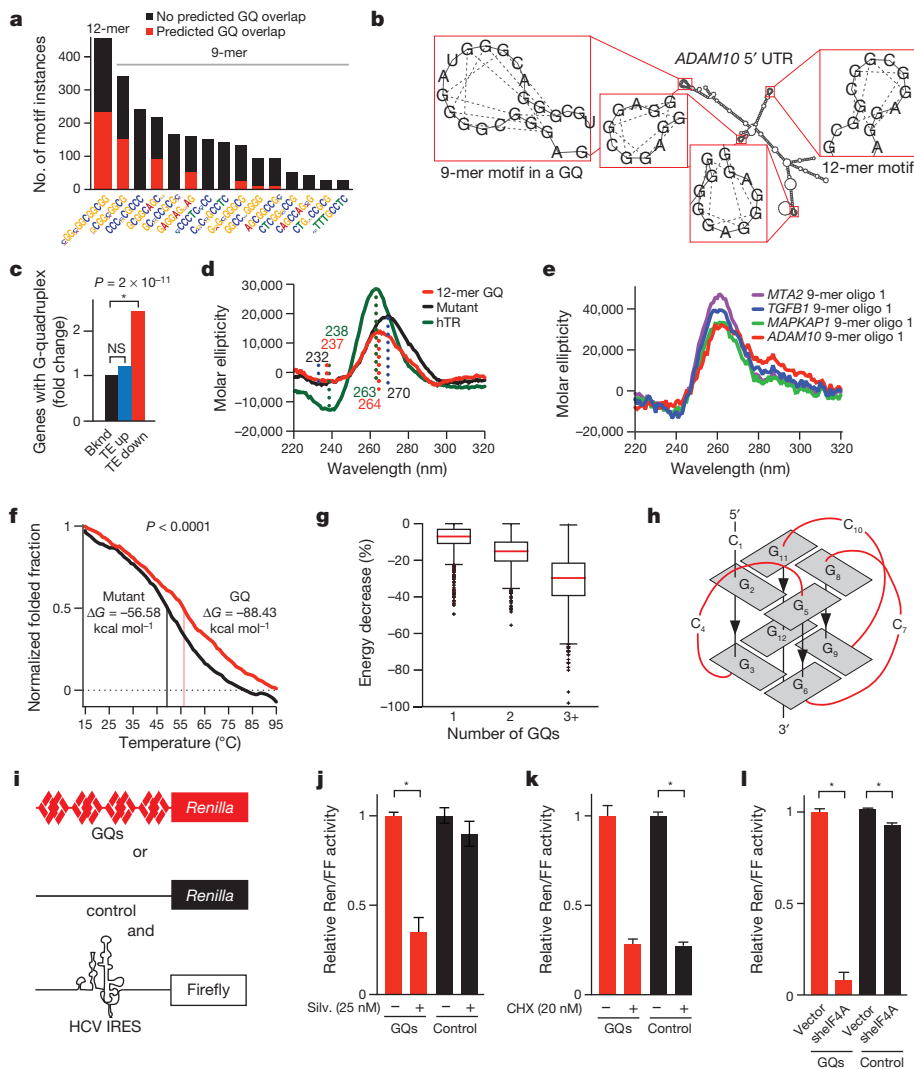


Figure 4 | G-quadruplex structures confer eIF4A-dependent translation. **a**, Bar graph indicating the motif prevalence and likelihood of forming G-quadruplexes (GQs; red). **b**, The ADAM10 5' UTR illustrates 12-nucleotide and 9-nucleotide motifs and G-quadruplexes. **c**, Enrichment of predicted 5' UTR G-quadruplex structures in the TE down gene set. Bknd, background. **d**, Circular dichroism spectra scan of the 12-nucleotide motif (CGG)₄ mutant oligomer (equal length and GC content) and human telomeric RNA (hTR), with known G-quadruplex structure folded in KCl. $n = 5$ replicates. **e**, Circular dichroism spectra scan of 9-nucleotide motifs with a 2-nucleotide flank from the 5' UTR of indicated genes folded with KCl. $n = 5$ replicates. **f**, Melting curve for circular dichroism spectra scan at a wavelength of 264 nm for the 12-nucleotide sequence (CGG)₄ and mutant oligomer. ΔG , free energy of unfolding. **g**, Calculated decrease in free energy for cellular UTRs with 1, 2 or 3+ motifs when allowed to fold into G-quadruplex structures. Red lines, medians; black lines, quartile boundaries. **h**, Diagram of parallel G-quadruplex conformation. Subscript numbers indicate nucleotide position. **i**, Schematic of reporter constructs with four 12-nucleotide motifs (G-quadruplexes, red), random sequence matched for length and GC content (control, black), and HCV IRES (white). **j**, **k**, Relative Renilla luciferase (Ren; normalized to IRES-firefly (FF)) expressed from the G-quadruplex (red) or control construct (black), treated with silvestrol (Silv.; j) or cycloheximide (CHX; k) for 24 h. $*P < 0.05$, $n = 3$ biological replicates and $n = 2$ technical replicates. **l**, Assay as in j and k comparing empty vector and shIF4A. $*P < 0.05$, $n = 3$ biological replicates and $n = 2$ technical replicates.

ADAM10 (Supplementary Table 5). The human telomeric sequence, the 12-nucleotide motif, and the 9-nucleotide motifs showed typical positive and negative molar ellipticity peaks at 264 nm and 240 nm, indicating parallel G-quadruplex structures. By contrast, the random oligomer had a shift in peak wavelengths (270 and 233 nm) (Fig. 4d, e and Extended Data Fig. 6b). This effect was dependent on the presence of potassium and was abrogated in sodium phosphate buffer without potassium (Extended Data Fig. 6c). Circular dichroism combined with thermal unfolding revealed that under potassium conditions the melting temperature for the 12-nucleotide motif was higher (56 °C) than the random oligomer (49 °C), corresponding to a free energy difference of -32 kcal mol⁻¹ (Fig. 4f). Similarly, computational modelling of complete 5' UTRs with increasing numbers of motifs showed a decrease of the predicted minimum free energy when the 5' UTR is allowed to fold into G-quadruplexes (Fig. 4g). Hence, the 12-nucleotide and some of the 9-nucleotide motifs can form energetically favourable G-quadruplex structures (Fig. 4h).

Next, we tested the 12-nucleotide (CGG)₄ motif in a translation assay. We constructed a reporter system to compare four 12-nucleotide motifs in tandem (G-quadruplex construct) to a random sequence of equal length and GC content (control construct), using IRES firefly luciferase as an internal control (Fig. 4i). Treatment with silvestrol (25 nM) reduced the translation of the G-quadruplex construct and did not affect the control construct. Hippuristanol and pateamine A had identical effects (Extended Data Fig. 6d), whereas cycloheximide (20 nM) suppressed both reporters (Fig. 4j, k). Other RNA helicases (DHX9, DHX36) may

resolve G-quadruplex structures^{30,31}. However, we find predominant expression of eIF4A2 in T-ALL³², and knockdown confirms an eIF4A-dependent effect on the G-quadruplex reporter (Fig. 4l and Extended Data Fig. 6e). Conversely, increased expression of eIF4A or a silvestrol-binding-site mutant of eIF4A (P159Q)³³ renders the G-quadruplex reporter insensitive to silvestrol (Extended Data Fig. 6f). Hence, eIF4A is limiting for the translation of mRNAs harbouring G-quadruplexes in their 5' UTRs.

Transcripts affected by silvestrol

The silvestrol-sensitive transcripts (TE down and rDiff gene lists) include many genes with known roles in T-ALL (Fig. 5a, b). The individual ribosome footprint distribution graphs (normalized for mean ribosome footprint count and gene length) illustrate recurrent patterns and also variations (Extended Data Fig. 7a). For example, the MYC, MDM2 and RUNX1 transcripts harbour multiple motifs in their 5' UTRs that correspond to peaks in ribosome footprint density, whereas house-keeping genes show no changes in ribosome footprint profiles. Gene ontology reveals a preponderance of transcription factors and oncogenes, but also some tumour suppressors (Extended Data Fig. 7b). Furthermore, we note a significant enrichment of superenhancer-associated genes³⁴—mostly transcription factors like NOTCH1, MYC, MYB, ETS1 and others (Fig. 5c, Extended Data Fig. 7c and Supplementary Table 6).

We confirmed the effects of silvestrol (25 nM for 24 h) on key target proteins (Fig. 5d and Extended Data Fig. 7a). The effect on the MYC protein was especially striking; it lasted for up to 48 h, and was readily

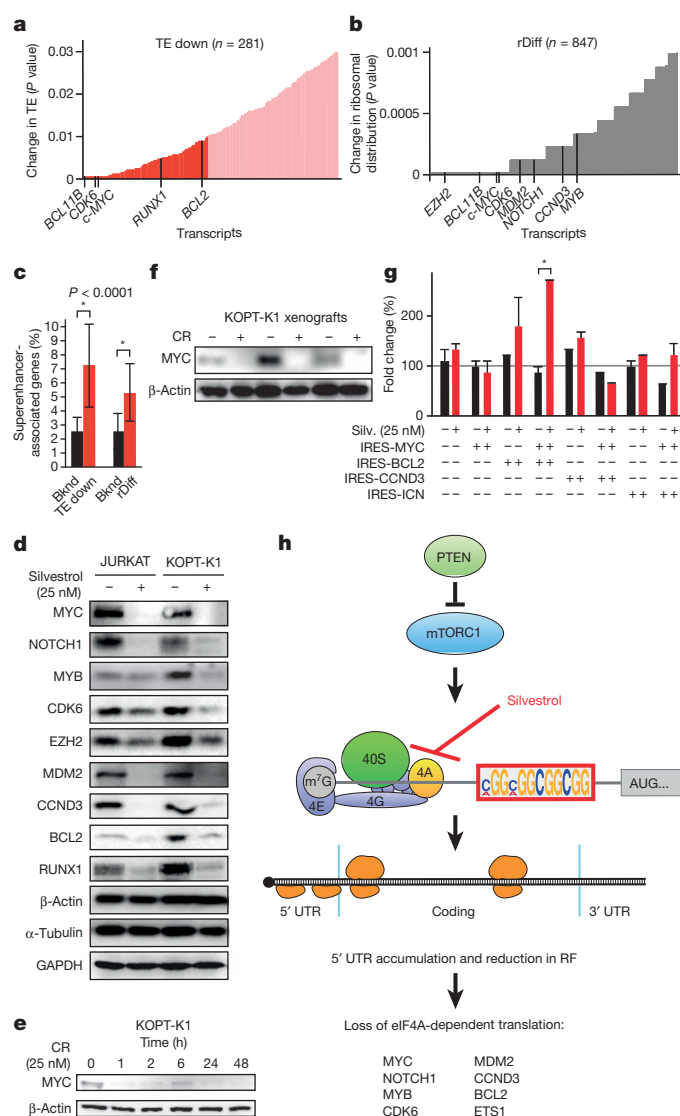


Figure 5 | Many oncogenes and transcription factors require eIF4A for translation. **a**, TE down genes ranked by TE (red, up to $P = 0.01$, see Fig. 3b). **b**, rDiff genes ranked by significance (up to $P = 0.001$, see Fig. 3c). **c**, Genes associated with superenhancers in T-ALL cells are enriched among TE down and rDiff gene sets. Bknd, background. $*P < 0.05$. Mean and s.d. are shown. **d**, Immunoblots of lysates from human T-ALL lines treated with silvestrol (25 nM, 24 h) and probed as indicated. **e**, Time-course analysis of protein expression in KOPT-K1 cells treated with (\pm)-CR-31-B (CR; 25 nM) for the indicated number of hours. **f**, Immunoblot on (\pm)-CR-31-B- or vehicle-treated KOPT-K1 xenografts, probed as indicated. **g**, Competition experiment (as in Fig. 1c, d) showing the percentage of each starting GFP-positive population of murine T-ALL cells partially transduced with the indicated constructs and treated with silvestrol (Silv.). $*P < 0.05$, $n = 3$ biological replicates, error bars indicate standard deviation. **h**, Diagram showing an eIF4A-dependent mechanism of translational control. RF, ribosome footprint.

detected in xenografts *in vivo* (Fig. 5d–f and Extended Data Fig. 8a, b). The corresponding mRNAs were unchanged, and the increase in MYC mRNA may reflect a known auto-regulatory mechanism³⁵ (Extended Data Fig. 8c). To test which proteins account for silvestrol's effects, we expressed several candidate genes and green fluorescent protein (GFP) in an eIF4A-independent manner from the HCV IRES. In mixed populations of transduced and parental murine leukaemic cells (see Fig. 1c), we found that individually genes were not protective, whereas co-expression of IRES-MYC and IRES-BCL2 afforded significant protection (Fig. 5g and Extended Data Fig. 8d–g). This result indicates that silvestrol kills cancer cells by disrupting the translation of several critical factors.

A selective mechanism of translational control

We report an eIF4A-dependent mechanism of translational control that is encoded in the 5' UTR of susceptible transcripts (Fig. 5h). The eIF4A RNA helicase has been implicated in the translation of mRNAs with long and complex 5' UTRs⁹. Ribosome footprinting technology allowed us to identify silvestrol-sensitive mRNAs and analyse their hallmark features. Besides 5' UTR length, these included a 12-nucleotide (CGG)₄ motif, additional 9-nucleotide motifs, and computationally predicted G-quadruplex structures that correspond to the motif. RNA G-quadruplexes are energetically favourable/stable RNA structures made from at least two stacks of four guanines with non-Watson–Crick interactions (Hoogsteen hydrogen bonds) with intervening linker nucleotides²⁹. The minimum number is 12 nucleotides—for example, the 12-nucleotide motif (CGG)₄—but shorter 9-nucleotide motifs can include neighbouring nucleotides to complete the structure. Frequently, more than 12 nucleotides contribute to G-quadruplexes, and typically silvestrol-sensitive transcripts harbour several G-quadruplexes in their 5' UTRs. These features define a specific subset of genes whose translation requires eIF4A and that is clearly distinct from mTORC1-dependent translation^{15,16}.

Many transcription factors and oncogenes require eIF4A for translation. G-quadruplex structures have been observed, for example, in the *NRAS* and *VEGFA* mRNAs^{36–38}, and computational predictions have suggested a broader role in translational regulation³⁹. Our results identify a broad set of eIF4A-sensitive transcripts including many oncogenes and transcriptional regulators (for example, MYC, MYB, NOTCH, CDK6, BCL2, and others). We can speculate that this mechanism provides additional control over these genes. In this regard, we note that a recent study identified a role for eIF4A2 in microRNA function and reports that eIF4A-sensitive transcripts are probably controlled by microRNAs¹³. We further find a significant overlap with genes whose transcription is controlled by superenhancers³⁴.

Our study provides new insight into the striking anticancer effect of silvestrol and related compounds. Several strategies are available to target oncoprotein translation (for example, mTORC inhibitors¹⁵, the eIF4G inhibitory peptide 4EGI-1 (ref. 40), or ribavirin⁴¹). We find that eIF4A has an oncogenic activity that is required for the expression of key oncogenes. Accordingly, silvestrol can obliterate oncoproteins including MYC, NOTCH, BCL2, and others. This finding suggests potential activity in other cancers as a single agent or in combination with other drugs.

METHODS SUMMARY

Ribosome footprinting. KOPT-K1 cells were treated with silvestrol (25 nM) or DMSO for 45 min, followed by cycloheximide treatment for 10 min. Total RNA and ribosome-protected fragments were isolated following a published protocol¹⁴. Deep sequencing libraries were generated from these fragments and sequenced on the HiSeq 2000 platform. Genome annotation was from the GENCODE project (<http://www.genencodegenes.org/releases/14.html>). Ribosome footprint intensity (RPM) and the expression value (RPKM) were measured from total mRNA-sequencing data and translation values were measured from ribosome footprint data. To evaluate the TE change between silvestrol- and vehicle-treated samples, we then calculated $TE = RPKM_{\text{footprint}}/RPKM_{\text{mRNA}}$ (ref. 15). Changes in ribosome footprint profiles were determined using the DEXseq algorithm²⁴.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 December 2013; accepted 1 May 2014.

Published online 27 July 2014.

- Wendel, H. G. *et al.* Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. *Nature* **428**, 332–337 (2004).
- Mamane, Y., Petroulakis, E., LeBacquer, O. & Sonenberg, N. mTOR, translation initiation and cancer. *Oncogene* **25**, 6416–6422 (2006).
- Rajasekhar, V. K. *et al.* Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol. Cell* **12**, 889–901 (2003).

4. Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
5. Lazaris-Karatzas, A., Montine, K. S. & Sonenberg, N. Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature* **345**, 544–547 (1990).
6. Ruggero, D. *et al.* The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.* **10**, 484–486 (2004).
7. Wendel, H. G. *et al.* Dissecting eIF4E action in tumorigenesis. *Genes Dev.* **21**, 3232–3237 (2007).
8. Mavrakis, K. J. *et al.* Tumorigenic activity and therapeutic inhibition of Rheb GTPase. *Genes Dev.* **22**, 2178–2188 (2008).
9. Parsyan, A. *et al.* mRNA helicases: the tacticians of translational control. *Nature Rev. Mol. Cell Biol.* **12**, 235–245 (2011).
10. Schatz, J. H. *et al.* Targeting cap-dependent translation blocks converging survival signals by AKT and PIM kinases in lymphoma. *J. Exp. Med.* **208**, 1799–1807 (2011).
11. Bordeleau, M. E. *et al.* Therapeutic suppression of translation initiation modulates chemosensitivity in a mouse lymphoma model. *J. Clin. Invest.* **118**, 2651–2660 (2008).
12. Marintchev, A. *et al.* Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. *Cell* **136**, 447–460 (2009).
13. Meijer, H. A. *et al.* Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. *Science* **340**, 82–85 (2013).
14. Ingolia, N. T., Ghaemmamghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
15. Thoreen, C. C. *et al.* A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485**, 109–113 (2012).
16. Hsieh, A. C. *et al.* The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485**, 55–61 (2012).
17. Palomero, T. *et al.* Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. *Nature Med.* **13**, 1203–1210 (2007).
18. Weng, A. P. *et al.* c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. *Genes Dev.* **20**, 2096–2109 (2006).
19. Zenatti, P. P. *et al.* Oncogenic *IL7R* gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia. *Nature Genet.* **43**, 932–939 (2011).
20. Pear, W. S. *et al.* Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated *Notch* alleles. *J. Exp. Med.* **183**, 2283–2291 (1996).
21. Rong, L. *et al.* Control of eIF4E cellular localization by eIF4E-binding proteins, 4E-BPs. *RNA* **14**, 1318–1327 (2008).
22. Rodrigo, C. M., Cencic, R., Roche, S. P., Pelletier, J. & Porco, J. A. Synthesis of rocaglamide hydroxamates and related compounds as eukaryotic translation inhibitors: synthetic and biological studies. *J. Med. Chem.* **55**, 558–562 (2012).
23. Choo, A. Y., Yoon, S. O., Kim, S. G., Roux, P. P. & Blenis, J. Rapamycin differentially inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation. *Proc. Natl Acad. Sci. USA* **105**, 17414–17419 (2008).
24. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).
25. Drewe, P. *et al.* Accurate detection of differential RNA processing. *Nucleic Acids Res.* **41**, 5189–5198 (2013).
26. Meyuhas, O. Synthesis of the translational apparatus is regulated at the translational level. *Eur. J. Biochem.* **267**, 6321–6330 (2000).
27. Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320–325 (1988).
28. Bailey, T. L. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659 (2011).
29. Bugaut, A. & Balasubramanian, S. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.* **40**, 4727–4741 (2012).
30. Booy, E. P. *et al.* The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. *Nucleic Acids Res.* **40**, 4110–4124 (2012).
31. Chakraborty, P. & Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst.)* **10**, 654–665 (2011).
32. Van Vlierberghe, P. *et al.* ETV6 mutations in early immature human T cell leukemias. *J. Exp. Med.* **208**, 2571–2579 (2011).
33. Sadlish, H. *et al.* Evidence for a functionally relevant rocaglamide binding site on the eIF4A-RNA complex. *ACS Chem. Biol.* **8**, 1519–1527 (2013).
34. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
35. Penn, L. J., Brooks, M. W., Laufer, E. M. & Land, H. Negative autoregulation of c-myc transcription. *EMBO J.* **9**, 1113–1121 (1990).
36. Kumari, S., Bugaut, A., Huppert, J. L. & Balasubramanian, S. An RNA G-quadruplex in the 5' UTR of the *NRAS* proto-oncogene modulates translation. *Nature Chem. Biol.* **3**, 218–221 (2007).
37. Shahid, R., Bugaut, A. & Balasubramanian, S. The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression. *Biochemistry* **49**, 8300–8306 (2010).
38. Morris, M. J., Negishi, Y., Pazsint, C., Schonhoft, J. D. & Basu, S. An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *J. Am. Chem. Soc.* **132**, 17831–17839 (2010).
39. Huppert, J. L., Bugaut, A., Kumari, S. & Balasubramanian, S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* **36**, 6260–6268 (2008).
40. Moerke, N. J. *et al.* Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128**, 257–267 (2007).
41. Kentsis, A., Topisirovic, I., Culjkovic, B., Shao, L. & Borden, K. L. Ribavirin suppresses eIF4E-mediated oncogenic transformation by physical mimicry of the 7-methyl guanosine mRNA cap. *Proc. Natl Acad. Sci. USA* **101**, 18105–18110 (2004).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the members of A.L.W.'s thesis committee: N. Rosen, A. M. Brown and S. W. Lowe. For reagents and advice we thank J. T. Barata, W. S. Pear, R. Cencic, S. Shuman, J. Cools, A. A. Ferrando, C. S. Fraser, N. J. Lajkiewicz, A. Luz, J. F. Glickman, C. Y. Park, P. Yellen, A. Heguy, K. Huberman and A. Viale. H.-G.W. is a Scholar of the Leukemia and Lymphoma Society. This research was supported by National Cancer Institute R01-CA142798-01 (H.-G.W.), the Leukemia Research Foundation (H.-G.W.), the Experimental Therapeutics Center (H.-G.W.), the American Cancer Society 10284 (H.-G.W.), European Union grant no. PITN-GA-2012-316861 (Y.Z.), the Fund for Scientific Research FWO Flanders (J.V.d.M. and P.R.), grants G.0198.08 and G.0869.10N (F.S.), the GOA-Ugent 12051203 (F.S.), Stichting tegen Kanker (F.S.), the Belgian Program of Interuniversity Poles of Attraction (F.S.), the Belgian Foundation Against Cancer (F.S.), the American Cancer Society PF-11-077-01-CDD (C.M.R.), the Lymphoma Research Foundation (J.H.S.), National Institutes of Health grants GM-067041 and GM-073855 (J.A.P.), and the Canadian Institutes of Health Research MOP-10653 (J.P.).

Author Contributions A.L.W. performed *in vivo* and treatment studies; K.S. performed ribosome footprinting and RNA structure studies; Y.Z. and P.D. analysed footprint data; V.K.R., V.R.S., K.J.M., M.J., J.E.R., J.H.S., C.Z., J.T.-F. and M.A.K. contributed to experiments; C.M.R. prepared (±)-CR-31-B; E.d.S. directed murine drug toxicity experiments; J.V.d.M., P.R. and F.S. generated genomic data on T-ALL; J.A.P. Jr and J.P. advised on all aspects of the study; G.R. supervised computational analyses; H.-G.W. designed the study and wrote the paper.

Author Information The ribosome footprinting and total mRNA sequencing data have been deposited in the Gene Expression Omnibus database under accession number GSE56887. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.-G.W. (wendelh@mskcc.org) or G.R. (Ratschg@mskcc.org).

The Laniakea supercluster of galaxies

R. Brent Tully¹, H  l  ne Courtois², Yehuda Hoffman³ & Daniel Pomar  de⁴

Galaxies congregate in clusters and along filaments, and are missing from large regions referred to as voids. These structures are seen in maps derived from spectroscopic surveys^{1,2} that reveal networks of structure that are interconnected with no clear boundaries. Extended regions with a high concentration of galaxies are called ‘superclusters’, although this term is not precise. There is, however, another way to analyse the structure. If the distance to each galaxy from Earth is directly measured, then the peculiar velocity can be derived from the subtraction of the mean cosmic expansion, the product of distance times the Hubble constant, from observed velocity. The peculiar velocity is the line-of-sight departure from the cosmic expansion and arises from gravitational perturbations; a map of peculiar velocities can be translated into a map of the distribution of matter³. Here we report a map of structure made using a catalogue of peculiar velocities. We find locations where peculiar velocity flows diverge, as water does at watershed divides, and we trace the surface of divergent points that surrounds us. Within the volume enclosed by this surface, the motions of galaxies are inward after removal of the mean cosmic expansion and long range flows. We define a supercluster to be the volume within such a surface, and so we are defining the extent of our home supercluster, which we call Laniakea.

The distribution of matter can be determined by two independent methods: either based on surveys of the distribution of galaxies in projection and redshift, or from the motions of galaxies. With the former, using galaxy redshift surveys, the assumption is required that the galaxy ‘lighthouses’ and mass distribution are strongly correlated, a condition that requires confirmation if, as is suspected, only a minor fraction of matter is baryonic. Moreover, with the former there is a stringent demand that the survey be complete, or at least that its incompleteness be well understood. With the latter, studies of galaxy motions, sparse sampling is acceptable (indeed inevitable) but dealing with errors is a challenge. Except for the very closest galaxies, uncertainties in distance measurements translate into uncertainties in the peculiar velocities of galaxies that are larger in amplitude than the actual peculiar velocities. Many measurements are required for suitable averaging and care must be taken to avoid systematic errors. Overall, the two paths to determining the distribution of matter are in good agreement, a consequence that represents a considerable success for the standard model of structure formation via gravitational instability^{4–7}.

The path from velocities to mass distributions benefits from the coherence in velocities on large scales. Multipole components in the velocity field can point to tidal influences beyond the survey region. The current all-sky redshift surveys and distance measurement surveys reach similar depths, but the latter probe structure to greater distances because of sensitivity to uncharted attractors and repellers. Coherence in motions on large scales means that signals can be measured by averaging over data in circumstances where individual contributions are very noisy.

Details about the actual measurement of galaxy distances and the derivation of peculiar velocities are given in Methods. These two parameters are available for more than 8,000 galaxies, affording extremely detailed information locally, degrading outward to increasingly coarse coverage. We use this material to reconstruct the large-scale structure of the nearby Universe⁷.

We obtain the underlying three-dimensional velocity and density fields by the Wiener filter algorithm^{8,9}, assuming the standard model of cosmology as a Bayesian prior. Large-scale structure is assumed to develop from gravitational instabilities out of primordial random Gaussian fluctuations. The developing density and velocity fields retain their Gaussian properties as long as the growth is in the linear regime. It has been shown⁸ that with a random Gaussian field, the optimal Bayesian estimator of the field given the data is the Wiener filter minimal variance estimator. At the present epoch, large-scale structure has become nonlinear on small scales. However, it is an attractive feature of the velocity field that the break from linearity is only on scales of a few megaparsecs, an order of magnitude smaller in scale than the deviations from linearity for the density field. In any event, the present discussion concerns structure on scales of tens to hundreds of megaparsecs, comfortably in the linear regime.

The Wiener filter result is determined by the ratio of power to power + noise. Hence, the large-scale structure is strongly constrained nearby, where uncertainties are small and the coverage is extensive. At large distances, where the data become more sparse and noisy, the Wiener filter attenuates the recovered density and velocity fields to the null field that is expected in the absence of data. However in the linear regime there is coherence in galaxy flows on much larger scales than seen in density fluctuations. Tidal influences from beyond the surveyed regions can be manifested in cosmic flows on scales that exceed the coverage in measured distances by a factor of two (ref. 10).

The ultimate goal is to map the velocity field to a radius that completely encompasses the sources of the motion of the Local Group (of 631 km s^{-1}) reflected in the cosmic microwave background dipole¹¹. However, our knowledge of flows on large scales remains inevitably modulated by the extent of the data. Our analysis of the data in the Cosmicflows-2 galaxy catalogue (see Methods for details) tells us that a coherent flow extends across the full extent of the region that we can map, reaching the Shapley concentration¹². It is clear that we do not yet have a sufficiently extensive compendium of distances to bound the full source of our deviant motion from the cosmic expansion.

For the present discussion, we focus on intermediate-scale flow patterns. The standard model of cosmology predicts that on the scale we are considering the flow is irrotational: namely, the velocity field \mathbf{v} is the gradient of a potential ϕ , $\mathbf{v} = -\nabla\phi$. The local minima and maxima of the potential (attractors and repellers respectively) are the drivers of the large-scale flow. We can define a ‘basin of attraction’ as the volume containing all points whose flow lines converge at a given attractor. The large-scale structure can be characterized on scales of a few megaparsecs and above by attractors and their basins of attraction.

The Wiener filter provides a straightforward way of decomposing the velocity field into a local component that is induced by the distribution within a zone and a tidal residual¹³. In the linear regime, the velocity and density fields, \mathbf{v} and δ , are directly related: $\nabla \cdot \mathbf{v} = -H_0(f\Omega_m, \Omega_\Lambda)\delta$ where f depend on the cosmological matter and vacuum energy densities characterized by Ω_m and Ω_Λ . We specify a centre and radius, and the density field within the defined volume is assumed to be a cut-out from the full Wiener filter density map. Then the Poisson-like equation between velocity and density can be solved to derive the velocity

¹Institute for Astronomy, University of Hawaii, Honolulu, Hawaii 96822, USA. ²Université Claude Bernard Lyon I, Institut de Physique Nucléaire, Université Lyon I, CNRS/IN2P3, Lyon 69622, France. ³Rachael Institute of Physics, Hebrew University, Jerusalem 91904, Israel. ⁴Institut de Recherche sur les Lois Fondamentales de l'Univers, CEA/Saclay, 91191 Gif-sur-Yvette, France.

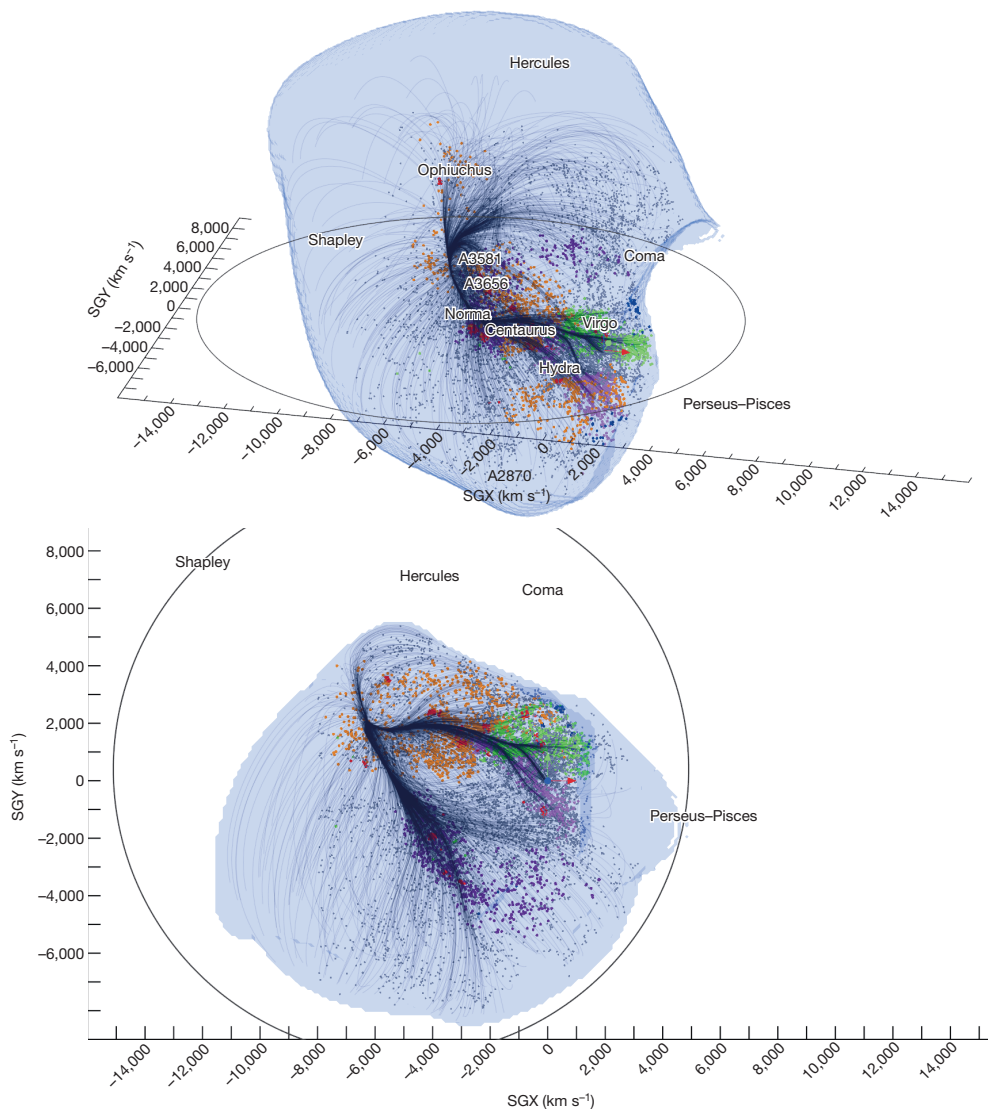


Figure 1 | Two views of the Laniakea supercluster. The outer surface (blue) demarcates the limits of local velocity flows. The plot is in supergalactic coordinates with axes SGX and SGY shown in the plane at SGZ = 0. Our Milky Way galaxy is at the origin. Units of $1,000 \text{ km s}^{-1}$ in velocity correspond to roughly 13.3 Mpc. Velocity streamlines are shown in black and terminate in the vicinity of the Norma cluster. Individual galaxies from a redshift catalogue are given colours to distinguish major components within the Laniakea supercluster: the historical Local supercluster in green, the Great Attractor region in orange, the Pavo-Indus filament in purple, and structures including the Antlia wall and Fornax-Eridanus cloud in magenta. Several major entities are named. Norma, Hydra, Centaurus, Virgo, Ophiuchus, A2870, A3581 and A3656 are individual clusters of galaxies embedded within the Laniakea supercluster. Shapley, Hercules, Coma and Perseus-Pisces are complexes of galaxies outside Laniakea. The outer black circle defines the domain used to separate between local and tidal flows. The panels provide two perspectives of the same scene.

field responding to just the matter within the prescribed volume. The vector subtraction of the local velocity component from the full flow gives the external component of the velocity field. The residual component is responsible for the bulk motion of the zone under consideration and for a quadrupole component within the zone. The decomposition allows us to probe the local velocity field, with the tidal field induced by distant structures filtered out. Relative attractors and their basins of attraction are defined with respect to that local field.

One more useful tool is now mentioned before we turn to results. At each position in space, the three eigenvalues of the velocity shear tensor can be calculated. If these eigenvalues are ordered from most positive to most negative, then a threshold can be set that captures four possibilities. Flows can be inward on all three axes, the condition of a cluster, inward on two axes and outward on the third, the condition of a filament, inward on one axis and outward on two, hence a sheet, or outward on all three axes, hence a void. Boundaries can be created around contiguous regions with the same shear properties, and the contours outline the cosmic web as reconstructed by the V-web algorithm¹⁴.

We now consider the full presentation of our results given in the Supplementary Video and in Extended Data Figs 1–5. We pay attention in turn to smaller (but still in the linear regime) scales to examine the separation of local and tidal flows and to isolate local basins of attraction. Particular attention has been given to locations where there are local divergences. In co-moving coordinates and with the removal of long-range

flows, there are places where relatively neighbouring galaxies can be found to be moving in opposing directions towards separate local basins of attraction. Voids are usually the demarcations between attraction basins, but divergences can occur along filaments and sheets. We note the very nearby case between our home basin of attraction and the Perseus-Pisces complex^{15,16}, where we find a particularly impressive example of velocity flows in an apparent bridge between attractors that are diverging. This divergence occurs in the feature called the Arch in Extended Data Fig 2. Similar structures abound on close inspection. Velocity information reveals the locations of divergence along filaments between high density regions. This dissipation of the cosmic web is expected with the accelerated expansion of the Universe. We emphasize that peculiar velocity information can reveal details that are otherwise hard to discern.

The particular interest in the present discussion is with the largest structure that can be circumscribed within the currently available distance and peculiar-velocity data: the structure schematically illustrated in Figs 1 and 2. The region includes 13 Abell clusters (with the Virgo cluster). Local flows within the region converge towards the Norma and Centaurus clusters, in good approximation to the location of what has been called the ‘Great Attractor’¹⁷. This volume includes the historical Local and Southern superclusters¹⁸, the important Pavo-Indus filament, an extension to the Ophiuchus cluster, the Local Void, and the Sculptor and other bounding voids. This region of inflow towards a local basin

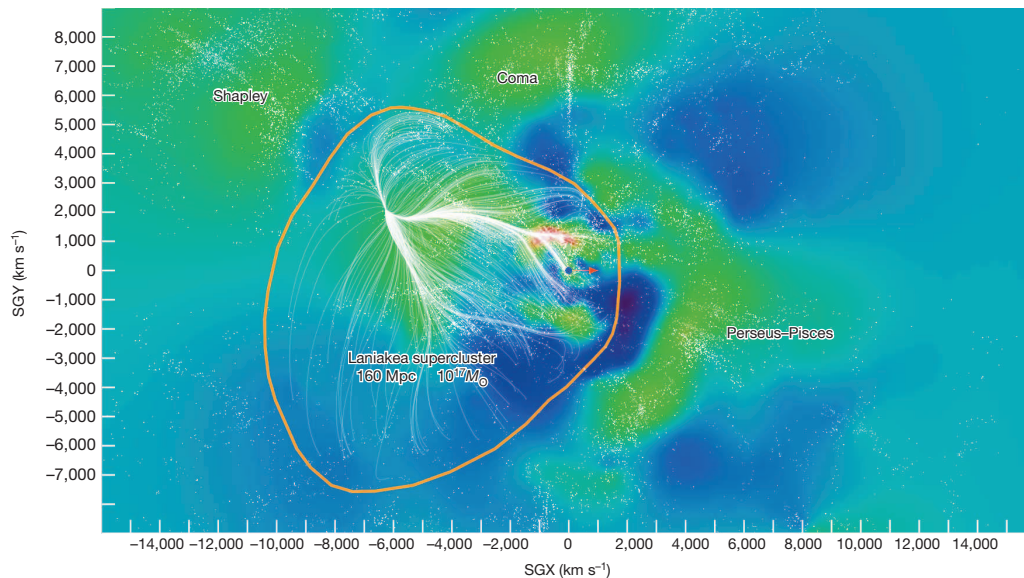


Figure 2 | A slice of the Laniakea supercluster in the supergalactic equatorial plane. Shaded contours represent density values within the equatorial slice, with red at high densities, green at intermediate densities and blue in voids. Our Milky Way galaxy is located at the black dot at the origin of the supergalactic coordinates system: a red arrow points right from the black dot toward increasing SGX and a green arrow points up toward increasing SGY. Individual galaxies from a redshift catalogue are given as white dots. Velocity flow streams within the Laniakea basin of attraction are shown in white. The orange contour encloses the outer limits of these streams. This domain has an extent of $\sim 12,000 \text{ km s}^{-1}$ ($\sim 160 \text{ Mpc}$ diameter) and encloses $\sim 10^{17}$ solar masses, M_{\odot} .

of attraction can be reasonably called a supercluster. The region, if approximated as round, has a diameter of $12,000 \text{ km s}^{-1}$ in units of the cosmic expansion or 160 megaparsecs, and encompasses $\sim 10^{17}$ solar masses. We propose that this region be named the Laniakea supercluster of galaxies (from the Hawaiian; *lani*, heaven, and *akea*, spacious, immeasurable).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 April; accepted 7 July 2014.

- de Lapparent, V., Geller, M. J. & Huchra, J. P. A slice of the universe. *Astrophys. J.* **302**, L1–L5 (1986).
- Gott, J. R. III *et al.* A map of the universe. *Astrophys. J.* **624**, 463–484 (2005).
- Courtois, H. M. *et al.* Cosmography of the local universe. *Astron. J.* **146**, 69 (2013).
- Strauss, M. A. *et al.* A redshift survey of IRAS galaxies. IV — The galaxy distribution and the inferred density field. *Astrophys. J.* **385**, 421–444 (1992).
- Dekel, A. *et al.* IRAS galaxies versus POTENT mass: density fields, biasing, and omega. *Astrophys. J.* **412**, 1–21 (1993).
- Kitaura, F.-S. *et al.* Cosmic structure and dynamics of the local Universe. *Mon. Not. R. Astron. Soc.* **427**, L35–L39 (2012).
- Courtois, H. M. *et al.* Three-dimensional velocity and density reconstructions of the local universe with Cosmicflows-1. *Astrophys. J.* **744**, 43 (2012).
- Zaroubi, S. *et al.* Wiener reconstruction of the large-scale structure. *Astrophys. J.* **449**, 446–459 (1995).
- Zaroubi, S., Hoffman, Y. & Dekel, A. Wiener reconstruction of large-scale structure from peculiar velocities. *Astrophys. J.* **520**, 413–425 (1999).
- Doumler, T. *et al.* Reconstructing cosmological initial conditions from galaxy peculiar velocities — II. The effect of observational errors. *Mon. Not. R. Astron. Soc.* **430**, 902–911 (2013).
- Fixsen, D. J. *et al.* The cosmic microwave background spectrum from the full COBE FIRAS data set. *Astrophys. J.* **473**, 576–587 (1996).
- Raychaudhury, S. The distribution of galaxies in the direction of the ‘Great Attractor’. *Nature* **342**, 251–255 (1989).
- Hoffman, Y. *et al.* The large-scale tidal velocity field. Preprint at <http://arXiv.org/abs/astro-ph/0102190> (2001).

- Hoffman, Y. *et al.* A kinematic classification of the cosmic web. *Mon. Not. R. Astron. Soc.* **425**, 2049–2057 (2012).
- Haynes, M. P. & Giovanelli, R. in *Large-Scale Motions in the Universe: A Vatican Study Week* (eds Rubin, V. C. & Coyne, G. V.) 31–70 (Princeton Univ. Press, 1988).
- Dekel, A. *et al.* POTENT reconstruction from mark III velocities. *Astrophys. J.* **522**, 1–38 (1999).
- Dressler, A. *et al.* Spectroscopy and photometry of elliptical galaxies — a large-scale streaming motion in the local universe. *Astrophys. J.* **313**, L37–L42 (1987).
- de Vaucouleurs, G. Evidence for a local supergalaxy. *Astron. J.* **58**, 30–32 (1953).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank our many collaborators in the accumulation of Cosmicflows-2 distances. We thank the CLUES collaboration, and in particular S. Gottlöber and J. Sorce in connection with the analysis. T. Jarrett provided an unpublished 2MASS Extended Source Catalog redshift compendium, the only all-sky redshift catalogue extensive enough to match the region of our reconstruction. The narration in the Supplementary Video is by S. Anvar and the original music is played by N.-E. Pomarède. The name ‘Laniakea’ was suggested by N. Napoleon, Kapiolani Community College, Hawaii. Financial support was provided by US National Science Foundation award AST09-08846, several awards through the Space Telescope Science Institute for observing time with Hubble Space Telescope, an award from the Jet Propulsion Lab for observations with Spitzer Space Telescope, and NASA award NNX12AE70G for analysis of data from the Wide-field Infrared Survey Explorer. We also acknowledge support from the Israel Science Foundation (1013/12) and the Lyon Institute of Origins under grant ANR-10-LABX-66 and the CNRS under PICS-06233.

Author Contributions R.B.T. guided the project, was involved in the data acquisition, interacted closely in the development of the ideas that are presented here, and wrote most of the Letter. H.C. was involved in the observing programme, was instrumental in coordinating activities, and was involved in all facets. Y.H. took responsibility for the theoretical analysis, including the Wiener filter, the cosmic web and the Malmquist bias correction. D.P. developed and applied visualization tools useful to this research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.B.T. (tully@ifa.hawaii.edu).

A 400-solar-mass black hole in the galaxy M82

Dheeraj R. Pasham^{1,2}, Tod E. Strohmayer² & Richard F. Mushotzky¹

M82 X-1, the brightest X-ray source in the galaxy M82, has been thought to be an intermediate-mass black hole (100 to 10,000 solar masses) because of its extremely high luminosity and variability characteristics^{1–6}, although some models suggest that its mass may be only about 20 solar masses^{3,7}. The previous mass estimates were based on scaling relations that use low-frequency characteristic time-scales which have large intrinsic uncertainties^{8,9}. For stellar-mass black holes, we know that the high-frequency quasi-periodic oscillations (100–450 hertz) in the X-ray emission that occur in a 3:2 frequency ratio are stable and scale in frequency inversely with black hole mass with a reasonably small dispersion^{10–15}. The discovery of such stable oscillations thus potentially offers an alternative and less ambiguous means of mass determination for intermediate-mass black holes, but

has hitherto not been realized. Here we report stable, twin-peak (3:2 frequency ratio) X-ray quasi-periodic oscillations from M82 X-1 at frequencies of 3.32 ± 0.06 hertz and 5.07 ± 0.06 hertz. Assuming that we can extrapolate the inverse-mass scaling that holds for stellar-mass black holes, we estimate the black hole mass of M82 X-1 to be 428 ± 105 solar masses. In addition, we can estimate the mass using the relativistic precession model, from which we get a value of 415 ± 63 solar masses.

Oscillations arising from general relativistic effects should scale inversely with black hole mass if they arise from orbital motion near the innermost stable circular orbit in the accretion disk^{16,17}, and there is observational support that they do for stellar-mass black holes¹⁰ (3–50 solar masses (M_{\odot})). Previous estimates of the mass of M82 X-1, of a few

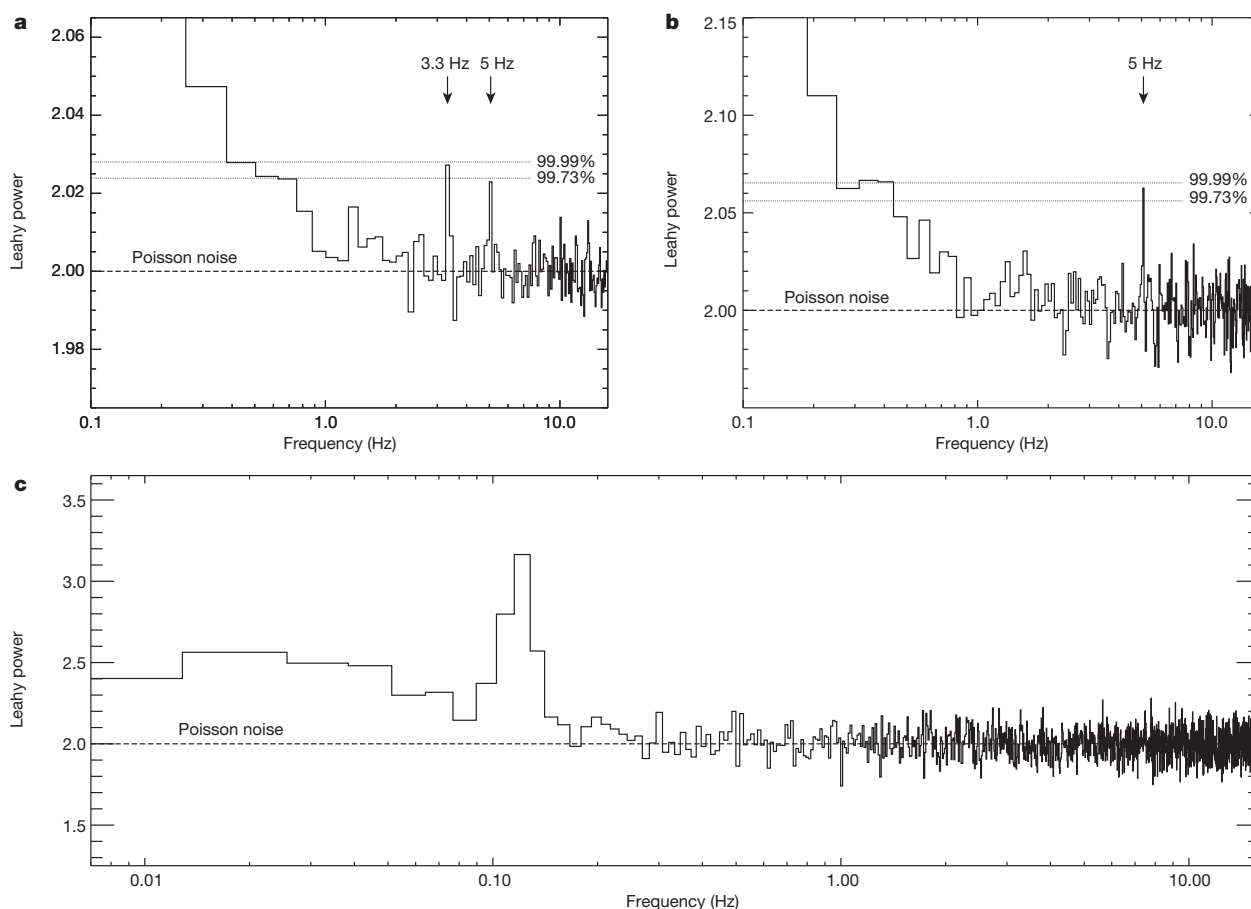


Figure 1 | Power density spectra of M82. **a**, Six-year average X-ray (3–13 keV) power density spectrum of M82 using 7,362 individual 128 s light curves. The frequency resolution is 0.125 Hz. The two strongest features in the power spectrum occur at 3.32 ± 0.06 and 5.07 ± 0.06 Hz, consistent with a 3:2 frequency ratio. **b**, Averaged power density spectrum of all (363) 1,024 s segments. The frequency resolution is 0.0625 Hz. The strongest feature is at

5 Hz. **c**, For a direct comparison with stellar-mass black holes, we show the broadband power density spectrum of M82 (using ~ 100 ks of XMM-Newton/EPIC data; ID: 0206080101) showing the low-frequency quasi-periodic oscillation at 120 mHz in addition to the high-frequency quasi-periodic oscillation pair in **a** and **b**. Power spectra are Leahy-normalized to a value of 2.

¹Astronomy Department, University of Maryland, College Park, Maryland 20742, USA. ²Astrophysics Science Division and Joint Space-Science Institute, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

hundred solar masses, combined with the type-C identification^{2,4,9} of its millihertz X-ray quasi-periodic oscillations, suggest that 3:2-ratio, twin-peak, high-frequency oscillations analogous to those seen in stellar-mass black holes, if present, should be detectable in the frequency range of a few hertz¹⁶. We accordingly searched NASA's Rossi X-ray Timing Explorer proportional counter array (PCA) archival data to look for 3:2 ratio oscillation pairs in the frequency range of 1–16 Hz, which corresponds to a black hole mass range of $50M_{\odot}$ – $2,000M_{\odot}$.

We detected two power spectral peaks at 3.32 ± 0.06 Hz (with a coherence, the ratio of frequency to width, of $Q = \nu/\Delta\nu > 27$) and 5.07 ± 0.06 Hz ($Q > 40$), consistent with a 3:2 frequency ratio (Fig. 1a, b). The combined statistical significance of the detection is greater than 4.7σ (Methods).

The PCA's field of view ($1^{\circ} \times 1^{\circ}$) of M82 includes a number of accreting X-ray sources in addition to M82 X-1¹⁸. The remarkable stability of the two quasi-periodic oscillations on timescales of a few years (Supplementary Videos 1 and 2), their 3:2 frequency ratio and their high oscillation luminosities strongly suggest they are not low-frequency quasi-periodic oscillations from a contaminating stellar-mass black hole (Methods). Also, a pulsar origin is very unlikely for several reasons. First, a pulsar signal would be much more coherent than that of the observed quasi-periodic oscillations, which have a finite width. Second, judging from the observed high quasi-periodic oscillation luminosities, it is extremely implausible that they originate from a pulsar (Methods). Finally, it would be highly unlikely to have two pulsars in the same field of view with spins in the 3:2 ratio. Also, on the basis of the average power spectrum of the background sky and a sample of accreting supermassive black holes monitored by the PCA in the same epoch as M82, we rule out an instrumental origin for these oscillations (Extended Data Figs 2–4). This leaves M82 X-1, which is persistently the brightest source in the field of view, as the most likely source associated with the 3:2-ratio quasi-periodic oscillation pair.

We estimated M82 X-1's black hole mass, assuming first the inverse-mass scaling of stellar-mass black holes and then the relativistic precession model^{19,20}, to be $(428 \pm 105)M_{\odot}$ in the first case and $(415 \pm 63)M_{\odot}$ in the second (Fig. 2). Combining the average 2–10 keV X-ray luminosity^{21,22} of the source, of $5 \times 10^{40} \text{ erg s}^{-1}$, with the measured mass suggests that the source is accreting close to the Eddington limit with an accretion efficiency of 0.8 ± 0.2 .

The previous mass measurements of M82 X-1 have large uncertainties owing to both systematic and measurement errors. For example, modelling of its X-ray energy spectra during the thermal-dominant state using a fully relativistic, 'multicoloured' disk model suggests that it hosts an intermediate-mass black hole that has a mass anywhere in the range of $200M_{\odot}$ – $800M_{\odot}$ and is accreting near the Eddington limit³. Also, the X-rays from this source are known to modulate with a periodicity of 62 days, which has been argued to be the orbital period of an intermediate-mass black hole²²—formed in the nearby star cluster MCG-11 by stellar runaway collisions^{5,23,24}—accreting matter via Roche lobe overflow from a $22M_{\odot}$ – $25M_{\odot}$ companion star²⁴. Detailed stellar binary evolution simulations suggest that the long periodicity is best explained by an intermediate-mass black hole with mass in the range of $200M_{\odot}$ – $5,000M_{\odot}$ (refs 5, 24). However, a recent study finds that this periodicity may instead be due to a precessing accretion disk, in which case a stellar-mass black hole will suffice to explain the apparent long periodicity²¹.

One of the main lines of argument for an intermediate-mass black hole in M82 X-1 follows from the assumption that its millihertz quasi-periodic oscillations are analogous to the type-C oscillations of stellar-mass black holes and to scale the inverse-mass relationship of the frequency of these millihertz oscillations (frequency range of 37–210 mHz (refs 9, 25); Fig. 1c) to the type-C, low-frequency X-ray oscillations (frequency range of 0.2–15 Hz (refs 10, 26)). There are two uncertainties with such

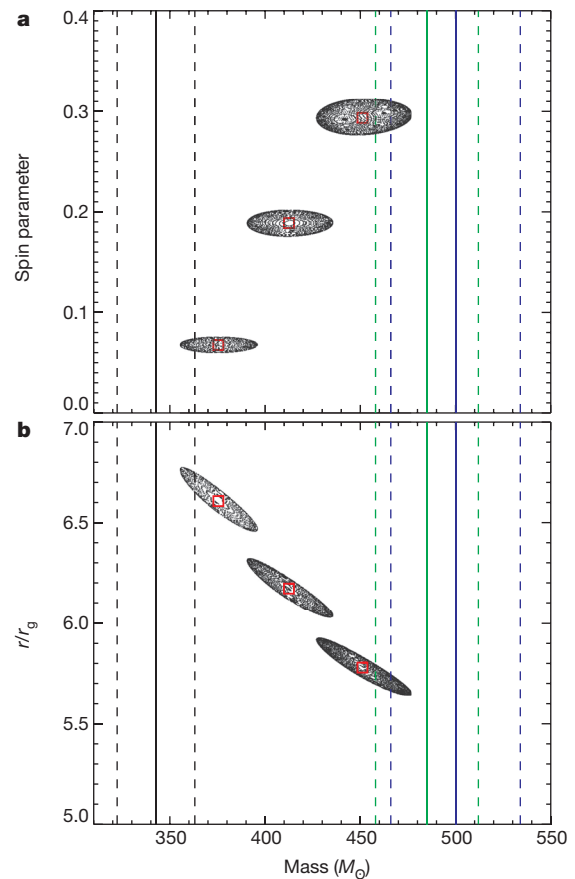


Figure 2 | Mass, spin and radius measurements. **a**, Contours (90% confidence) of M82 X-1's mass as a function of the spin parameter, that is, the ratio of the black hole's angular momentum to its mass. The three contours correspond to the three low frequency values (37, 120 and 210 mHz), with the mass increasing as the low-frequency oscillation frequency increases.

b, Contours of M82 X-1's mass as a function of the radius of the origin of these oscillations (in units of $r_g = GM/c^2$, where G , M and c are the gravitational constant, the black hole mass and the speed of light, respectively). In **a** and **b**, the vertical lines (solid, solution; dashed, upper and lower limits) represent M82 X-1's mass estimates assuming a simple inverse-mass scaling for the high-frequency quasi-periodic oscillations. The three colours correspond to scalings using the microquasars GRO J1655-40²⁸ (green), XTE J1550-64²⁹ (blue) and GRS 1915+105³⁰ (black).

scaling: first, it was unclear—until now—whether these millihertz oscillations are indeed the type-C analogues of stellar-mass black holes^{8,9}, and, second, both the type-C and the millihertz oscillations are variable, resulting in a large dispersion in the measured mass of $25M_{\odot}$ – $1,300M_{\odot}$ (refs 1, 2, 4, 6). The simultaneous discovery of stable, 3:2 ratio, high-frequency periodicities and low-frequency millihertz oscillations allows us to set the overall frequency scale of the X-ray power spectrum. This result not only implies that the millihertz quasi-periodic oscillations of M82 X-1 are the type-C analogues of stellar-mass black holes, but also provides an independent, and so far the most accurate, black hole mass measurement.

Finally, it should be pointed out that although the root mean squared amplitudes (3–5% of the X-ray flux) of the oscillations reported here, and their frequencies with respect to the millihertz oscillations (a few orders of magnitude higher), are similar to those observed in stellar-mass black holes¹⁰, they appear narrower (with Q values of 27, 40 and 80) than stellar-mass black holes, which have Q values of <20 (refs 14, 27). In stellar-mass black holes, the Q factor seems to be energy dependent in some cases¹⁴, and it is plausible that a similar effect may be operating in this case.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 April; accepted 21 July 2014.

Published online 17 August 2014.

- Casella, P. *et al.* Weighing the black holes in ultraluminous X-ray sources through timing. *Mon. Not. R. Astron. Soc.* **387**, 1707–1711 (2008).
- Dewangan, G. C., Titarchuk, L. & Griffiths, R. E. Black hole mass of the ultraluminous X-ray source M82 X-1. *Astrophys. J.* **637**, L21–L24 (2006).
- Feng, H. & Kaaret, P. Identification of the X-ray thermal dominant state in an ultraluminous X-ray source in M82. *Astrophys. J.* **712**, L169–L173 (2010).
- Mucciarelli, P., Casella, P., Belloni, T., Zampieri, L. & Ranalli, P. A variable quasi-periodic oscillation in M82 X-1. Timing and spectral analysis of XMM-Newton and RossiXTE observations. *Mon. Not. R. Astron. Soc.* **365**, 1123–1130 (2006).
- Patruno, A., Portegies Zwart, S., Dewi, J. & Hopman, C. The ultraluminous X-ray source in M82: an intermediate-mass black hole with a giant companion. *Mon. Not. R. Astron. Soc.* **370**, L6–L9 (2006).
- Zhou, X.-L., Zhang, S.-N., Wang, D.-X. & Zhu, L. Calibrating the correlation between black hole mass and X-ray variability amplitude: X-ray only black hole mass estimates for active galactic nuclei and ultra-luminous X-ray sources. *Astrophys. J.* **710**, 16–23 (2010).
- Okajima, T., Ebisawa, K. & Kawaguchi, T. A stellar-mass black hole in the ultraluminous X-ray source M82 X-1? *Astrophys. J.* **652**, L105–L108 (2006).
- Middleton, M. J., Roberts, T. P., Done, C. & Jackson, F. E. Challenging times: a re-analysis of NGC 5408 X-1. *Mon. Not. R. Astron. Soc.* **411**, 644–652 (2011).
- Pasham, D. R. & Strohmayer, T. E. On the nature of the mHz X-ray quasi-periodic oscillations from ultraluminous X-ray source M82 X-1: Search for timing-spectral correlations. *Astrophys. J.* **771**, 101 (2013).
- McClintock, J. E. & Remillard, R. A. in *Compact Stellar X-Ray Sources* (eds Lewin, W. & van der Klis, M.) 157–213 (Cambridge Univ. Press, 2006).
- Remillard, R. A., Morgan, E. H., McClintock, J. E., Bailyn, C. D. & Orosz, J. A. RXTE observations of 0.1–300 Hz quasi-periodic oscillations in the microquasar GRO J1655–40. *Astrophys. J.* **522**, 397–412 (1999).
- Remillard, R. A., Munro, M. P., McClintock, J. E. & Orosz, J. A. Evidence for harmonic relationships in the high-frequency quasi-periodic oscillations of XTE J1550–564 and GRO J1655–40. *Astrophys. J.* **580**, 1030–1042 (2002).
- Remillard, R. A., Sobczak, G. J., Munro, M. P. & McClintock, J. E. Characterizing the quasi-periodic oscillation behavior of the X-ray nova XTE J1550–564. *Astrophys. J.* **564**, 962–973 (2002).
- Strohmayer, T. E. Discovery of a 450 Hz quasi-periodic oscillation from the microquasar GRO J1655–40 with the Rossi X-Ray Timing Explorer. *Astrophys. J.* **552**, L49–L53 (2001).
- Strohmayer, T. E. Discovery of a second high-frequency quasi-periodic oscillation from the microquasar GRS 1915+105. *Astrophys. J.* **554**, L169–L172 (2001).
- Abramowicz, M. A., Kluźniak, W., McClintock, J. E. & Remillard, R. A. The importance of discovering a 3:2 twin-peak quasi-periodic oscillation in an ultraluminous X-ray source, or how to solve the puzzle of intermediate-mass black holes. *Astrophys. J.* **609**, L63–L65 (2004).
- Abramowicz, M. A. & Kluźniak, W. Interpreting black hole QPOs. Preprint at <http://arxiv.org/abs/astro-ph/0312396> (2003).
- Matsumoto, H. *et al.* Discovery of a luminous, variable, off-center source in the nucleus of M82 with the Chandra High-Resolution Camera. *Astrophys. J.* **547**, L25–L28 (2001).
- Stella, L., Vietri, M. & Morsink, S. M. Correlations in the quasi-periodic oscillation frequencies of low-mass X-ray binaries and the relativistic precession model. *Astrophys. J.* **524**, L63–L66 (1999).
- Motta, S. E., Belloni, T. M., Stella, L., Munoz-Darias, T. & Fender, R. Precise mass and spin measurements for a stellar-mass black hole through X-ray timing: the case of GRO J1655–40. *Mon. Not. R. Astron. Soc.* **437**, 2554–2565 (2014).
- Pasham, D. R. & Strohmayer, T. E. Can the 62 day X-ray period of ULX M82 X-1 be due to a precessing accretion disk? *Astrophys. J.* **774**, L16 (2013).
- Kaaret, P. & Feng, H. Confirmation of the 62 day X-ray periodicity from M82. *Astrophys. J.* **669**, 106–108 (2007).
- Voss, R., Nielsen, M. T. B., Nelemans, G., Fraser, M. & Smartt, S. J. On the association of ULXs with young superclusters: M82 X-1 and a new candidate in NGC 7479. *Mon. Not. R. Astron. Soc.* **418**, L124 (2011).
- Portegies Zwart, S. F., Baumgardt, H., Hut, P., Makino, J. & McMillan, S. L. W. Formation of massive black holes through runaway collisions in dense young star clusters. *Nature* **428**, 724–726 (2004).
- Caballero-García, M. D., Belloni, T. & Zampieri, L. Quasi-periodic oscillations and energy spectra from the two brightest ultra-luminous X-ray sources in M82. *Mon. Not. R. Astron. Soc.* **436**, 3262–3270 (2013).
- Casella, P., Belloni, T. & Stella, L. The ABC of low-frequency quasi-periodic oscillations in black hole candidates: analogies with Z sources. *Astrophys. J.* **629**, 403–407 (2005).
- Belloni, T. M., Sanna, A. & Mèndez, M. High-frequency quasi-periodic oscillations in black hole binaries. *Mon. Not. R. Astron. Soc.* **426**, 1701–1709 (2012).
- Beer, M. E. & Podsiadlowski, P. The quiescent light curve and the evolutionary state of GRO J1655–40. *Mon. Not. R. Astron. Soc.* **331**, 351–360 (2002).
- Orosz, J. A. *et al.* An improved dynamical model for the microquasar XTE J1550–564. *Astrophys. J.* **730**, 75–88 (2011).
- Steehgs, D. *et al.* The not-so-massive black hole in the microquasar GRS1915+105. *Astrophys. J.* **768**, 185–192 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work is based on observations made with the Rossi X-ray Timing Explorer, a mission that was managed and controlled by NASA's Goddard Space Flight Center in Greenbelt, Maryland, USA. All the data used in the present article is publicly available through NASA's HEASARC archive. D.R.P. would like to thank M. Tripp, C. Miller and C. Reynolds for discussions.

Author Contributions D.R.P. and T.E.S. reduced the data, carried out the analysis and wrote the paper. R.F.M. contributed to the interpretation of the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.R.P. (dheeraj@astro.umd.edu).

METHODS

Estimating the expected significance of the quasi-periodic oscillations. The detectability (statistical significance, n_σ) of a quasi-periodic oscillation feature can be expressed as

$$n_\sigma = \frac{1}{2} r^2 \frac{S^2}{S+B} \sqrt{\frac{T}{\Delta\nu}}$$

where r , S , B , T and $\Delta\nu$ are the root mean squared (r.m.s.) amplitude of the quasi-periodic oscillation, the source count rate, the background count rate, the exposure time and the width of the quasi-periodic oscillation, respectively³¹. Assuming an r.m.s. amplitude of a few per cent—similar to that seen in stellar-mass black holes²⁷—and using the mean Rossi X-ray Timing Explorer (RXTE)/PCA source and background rates obtained from prior observations²¹, we found that the wealth of publicly available, archival RXTE monitoring data ($\sim 10^6$ s) spread across a time span of ~ 6 years would suffice for a sensitive search for high-frequency quasi-periodic oscillations in M82 X-1.

Data primer. M82 was monitored (0.5–2 ks roughly once every three days) with the RXTE/PCA between 1997 February 2 and 1997 November 25 (0.8 years) and between 2004 September 2 and 2009 December 30 (5.3 years). All the PCA observations were carried out in the GoodXenon data acquisition mode. The total number of monitoring observations was 867, which were divided among six proposals (RXTE proposal IDs P20303, P90121, P90171, P92098, P93123 and P94123). As recommended by the data analysis guide provided by the RXTE Guest Observer Facility (<https://heasarc.gsfc.nasa.gov/docs/xte/abc/screening.html>), we first screened the data to include only time intervals that satisfy the following criteria: ELV > 10.0 && OFFSET < 0.02 && (TIME SINCE SAA < 0 || TIME SINCE SAA > 30) && ELECTRON2 < 0.1. In addition to the above standard filters, we used only X-ray events within the energy range of 3–13 keV, which translates to PCA channels 7–32. This energy range is comparable to the bandpass in which high-frequency quasi-periodic oscillations have been reported from stellar-mass black holes^{10,14,15}. Moreover, beyond 13 keV the background dominates the overall count rate by a factor greater than 10. For each observation, we used all active proportional counter units, to maximize the count rate and, thus, the sensitivity to quasi-periodic oscillations.

Before we extracted the individual power spectra, we created individual light curves (using a bin size of 1 s) from all the observations. Through manual inspection we removed a small number of observations affected by flares, because these are attributed to background events, for example gamma-ray bursts, not associated with the source. Extended Data Fig. 1 shows two sample light curves and their corresponding power spectra. They represent the typical quality of the individual light curves used in extracting the average power spectra.

Estimating the statistical significances. We first divided the data into 128 s segments and extracted their light curves with a time resolution of 1/32 s. We then constructed a Leahy-normalized³² power density spectrum (where the Poisson noise level is 2) from each of these 128 s light curves. All the power spectra were then combined (7,362 individual power spectra) to obtain a six-year-averaged power density spectrum of M82 (Fig. 1a).

To estimate the statistical significance of any features in the 1–16 Hz range of the six-year-averaged power density spectrum obtained using the 128 s data segments, we first ensured that the local mean was equal to 2, the value expected from a purely Poisson (white noise) process. We then computed the probability, at the 99.73% (3σ) and the 99.99% (3.9σ) confidence levels, of obtaining the power, $P = P_* \times 7,362 \times 16$ from a χ^2 distribution with $2 \times 7,362 \times 16$ degrees of freedom. Here P_* is the power value of a statistical fluctuation at a given confidence level. We used this χ^2 distribution because we summed 16 neighbouring frequency bins and divided by 16, and averaged 7,362 individual power spectra. Considering the total number of trials (frequency bins within 1–16 Hz), we computed the 99.73% (1/(371 trials)) and the 99.99% (1/(10,000 trials)) confidence limits (horizontal dotted lines in Fig. 1a). We detect two power spectral peaks at 3.32 ± 0.06 and 5.07 ± 0.06 Hz that are significant at the 2×10^{-4} (3.7σ) and 6×10^{-3} (2.75σ) levels, respectively, assuming both features were searched for independently between 1 and 16 Hz. However, after identifying the first feature at 3.3 Hz, if we are searching for a second feature at a 3:2 ratio, then the search from there on only includes the bins close to $3/2 \times 3.3$ Hz or $2/3 \times 3.3$ Hz. If this is taken into consideration, the significance of the 5 Hz feature increases, owing to the smaller number of trials, to 5×10^{-4} , or 3.5σ . To further test the significance of the 5 Hz feature, we extracted an average power density spectrum using all the data with segments longer than 1,024 s (Fig. 1b). The 5 Hz feature ($Q > 80$) is clearly detected at the 1.5×10^{-4} confidence level, or 3.8σ , considering a full search between 1 and 16 Hz.

The combined probability of two independent chance fluctuations in the 3:2 frequency ratio, one at the 3.7σ level (3.3 Hz feature) and the other at the 2.75σ level (5 Hz feature), is greater than 4.7σ .

Signal cannot be from a single observation. The very presence of these two features in the six-year-averaged power spectrum suggests that they are stable on this timescale. To further rule out the possibility that these oscillations are due to a single particular observation, or a small number of them, we constructed two dynamic average power spectra, one for the 128 s segments (dynamic power density spectrum#1; Supplementary Video 1) and another for the 1,024 s segments (dynamic power density spectrum#2; Supplementary Video 2). These track the evolution of the average power density spectra as a function of the total number of individual power spectra used in constructing the average. The two dynamic power density spectra clearly suggest that the power in these two features builds up gradually as more data are averaged, as opposed to appearing suddenly, which would be expected if a single or a small number of observations were contributing all the signal power. In addition, dynamic power density spectrum#1 clearly shows that while the 5 Hz feature is stronger during the earlier stages of the monitoring program, the 3.3 Hz feature is stronger during the later observations. Longer exposures of the order of 1–2 ks were carried out during the earlier stage of the monitoring program, which explains the higher significance of the 5 Hz feature in the average power density spectrum of the 1,024 s segments (Fig. 1b).

Root mean squared amplitude of the quasi-periodic oscillations. To calculate the r.m.s. variability amplitude of these quasi-periodic oscillations, we first determined the mean net count rate (source + background) of all the light curves used to extract the average power spectra. These values were equal to $29.9 \text{ counts s}^{-1}$ and $35.2 \text{ counts s}^{-1}$ for the power spectra with 128 s and 1,024 s segments, respectively. The r.m.s. amplitudes of the 3.3 and 5 Hz quasi-periodic oscillations, not correcting for the background, were estimated to be $1.1 \pm 0.1\%$ and $1.0 \pm 0.1\%$, respectively. Similarly, the r.m.s. amplitude of the 5 Hz feature in the power density spectrum with 1,024 s segments was estimated to be $1.1 \pm 0.1\%$. We then estimated the mean background count rate from the Standard2 data using the latest PCA background model. The mean background rates during the 128 s and the 1,024 s segments were estimated to be $18.9 \text{ counts s}^{-1}$ and $24.0 \text{ counts s}^{-1}$, respectively. After accounting for the X-ray background, we find that the r.m.s. amplitudes of the 3.3 and 5 Hz features—averaged over all of the data—are $3.0 \pm 0.4\%$ and $2.7 \pm 0.4\%$, respectively, and the amplitude of the 5 Hz feature within the 1,024 s power density spectrum was estimated to be $3.5 \pm 0.4\%$.

Furthermore, the source count rates estimated above (net minus background) represent the combined contribution from all the X-ray point sources within the PCA's $1^\circ \times 1^\circ$ field of view¹⁸. Thus, the quasi-periodic oscillation r.m.s. amplitudes are underestimated. A study using the high-resolution camera on board Chandra suggests that there are multiple point sources within the $1' \times 1'$ region around M82 X-1^{18,33}. Tracking the long-term variability of these sources suggests that the maximum luminosity reached by any of these sources—except for source 5 (as defined by ref. 18)—is less than one-fifth of the average luminosity of M82 X-1³³. Source 5 is a highly variable transient ULX with a 0.5–10 keV luminosity varying between 10^{37} and $10^{40} \text{ erg s}^{-1}$ (see the middle-left panel of fig. 1 of ref. 33). The quasi-periodic oscillations reported here are most probably produced by M82 X-1, which has persistently been the brightest source of any in the immediate vicinity of M82 X-1 (see the following sections). Although a precise value of the r.m.s. amplitude cannot be evaluated using the present data, we estimate an absolute upper limit by calculating the inverse of the fraction of the count rate contribution from M82 X-1, assuming all the remaining contaminating sources are at their brightest ever detected. This scenario is highly unlikely but will serve as an absolute upper bound on the r.m.s. amplitudes of the quasi-periodic oscillations, assuming they are from M82 X-1. Using the values reported by ref. 33, the fraction is roughly 1.8. Thus, the true r.m.s. amplitudes of the 3.3 and 5.5 Hz quasi-periodic oscillations are estimated to be in the range of 3–5%.

Also, XMM-Newton's EPIC instruments—with an effective area of about one-fifth of RXTE's PCA, albeit with lower background—observed M82 at multiple epochs, with a total effective exposure of ~ 350 ks. These observations were taken in the full-frame data acquisition mode with a time resolution of 73.4 ms or a Nyquist frequency of 6.82 Hz. This value is close to the quasi-periodic oscillation frequencies of interest and causes some signal suppression³¹. Nevertheless, we extracted an average 3–10 keV power density spectrum with 128 s data segments using all the observations (2,718 individual power spectra). We do not detect any statistically significant features near 3.3 and 5 Hz; however, we estimate quasi-periodic oscillation upper limits (3σ confidence) of 5.2 and 6.2% r.m.s. (using equations 4.4 and 4.10 of ref. 29) at 3.3 and 5 Hz, respectively, which are roughly twice the r.m.s. values of the quasi-periodic oscillations detected in the PCA data.

Using the 128 s data segments from RXTE, we also studied the energy dependence of the r.m.s. amplitudes of the two oscillations (Extended Data Table 1). Although the error bars are large, there seems to be a modest decrease in the r.m.s. amplitudes of these oscillations at lower X-ray energies. XMM-Newton's EPIC instruments are more sensitive in the 3–8 keV band, which is comparable to PCA channels 7–18 (the first and the fourth rows of Extended Data Table 1).

Ruling out low-frequency quasi-periodic oscillations. Low-frequency quasi-periodic oscillations of stellar-mass black holes, such as the type-C quasi-periodic oscillations²⁶, have typical centroid frequencies of a few hertz with r.m.s. amplitudes²⁶ of 5–25%, but are known to vary in frequency by factors of 8–10 over timescales of days^{34,35}. This would lead to very broad features in the kind of average power spectrum we have used. Moreover, among the plethora of low-frequency quasi-periodic oscillations at present known, there is no indication of them preferentially occurring with a 3:2 frequency ratio. Furthermore, the average luminosity of the quasi-periodic oscillations reported here is ~ 0.03 times the average luminosity of all the sources observed by the PCA in the 3–13 keV band^{21,22}, which is $\sim 0.03 \times 5 \times 10^{40} \text{ erg s}^{-1} = 1.5 \times 10^{39} \text{ erg s}^{-1}$. This is comparable to or more than the peak X-ray luminosities of the contaminating sources, except for source 5 (refs 33, 36). Therefore, if these features were simply low-frequency quasi-periodic oscillations produced by any of the contaminating sources—except for source 5—their X-ray flux would have to be modulated at almost 100%, which is not plausible for the typical amplitudes of low-frequency quasi-periodic oscillations.

Source 5, which is a ULX, could in principle be the origin of these 3:2 ratio quasi-periodic oscillations. However, 3–4 mHz quasi-periodic oscillations have been discovered from this source and have been identified as type-A/B quasi-periodic oscillations analogues of stellar-mass black holes³⁷. Such a characterization for the millihertz quasi-periodic oscillations suggests that the ULX might host a black hole with a mass of $12,000 M_{\odot}$ – $43,000 M_{\odot}$ (ref. 37). If that were the case, the expected frequency range of high-frequency quasi-periodic oscillation analogues for a few $10,000 M_{\odot}$ black hole would be a few 100s of millihertz, which is a factor of 10 lower than the quasi-periodic oscillations reported here, suggesting that the 3.3 and 5 Hz quasi-periodic oscillations are less likely to be the high-frequency quasi-periodic oscillation analogues of source 5.

Ruling out pulsar origin. Rotation-powered pulsars can be strongly excluded, because they simply cannot provide the required luminosity. A neutron star's rotational energy loss rate can be expressed in terms of its moment of inertia, I , spin period, P , and period derivative, \dot{P} , as

$$\dot{E}_{\text{rot}} = \frac{2\pi I \dot{P}}{P^3}$$

No known pulsar has a spin-down luminosity comparable to the estimated quasi-periodic oscillation X-ray luminosity. For example, the energetic Crab pulsar has $\dot{E} \approx 2 \times 10^{38} \text{ erg s}^{-1}$, and only a fraction of a pulsar's spin-down power typically appears as X-ray radiation. This rules out rotation-powered pulsars. Because M82 is a starburst galaxy, it probably hosts a population of accreting X-ray pulsars. Such accretion-powered pulsar systems are typically limited by the Eddington limit, of $\sim 2 \times 10^{38} \text{ erg s}^{-1}$, for a canonical neutron star. Useful comparisons can be made with the population observed with the RXTE/PCA in the Small Magellanic Cloud³⁸. The authors of ref. 38 presented pulsed luminosities for the Small Magellanic Cloud pulsar population, and none was larger than $\sim 3 \times 10^{38} \text{ erg s}^{-1}$. Again, this is much smaller than the inferred quasi-periodic oscillations luminosities. Moreover, such pulsars are variable, and their time-averaged luminosities would be reduced further by their outburst duty cycles. At present, the only pulsar that is known to reach luminosities of $\sim 10^{40} \text{ erg s}^{-1}$ for brief periods of time is GRO J1744-28—the 'bursting pulsar'^{39–41}. This object has a 2.1 Hz spin frequency and was discovered during an outburst that spanned the first 3 months of 1996 (we note that at the time of writing the source was detected in outburst again, suggesting a duty cycle of about 18 years^{42–47}). Its peak persistent luminosity (assuming a distance close to that of the Galactic centre) was $\sim 7 \times 10^{38} \text{ erg s}^{-1}$. With a pulsed amplitude of about 10%, this would still give a pulsed luminosity much less than the inferred luminosities of the quasi-periodic oscillations. The type-II—accretion driven—bursts from this source⁴⁸ could reach about $10^{40} \text{ erg s}^{-1}$, and with a 10% pulsed amplitude this could give an instantaneous luminosity close to that of the average quasi-periodic oscillation luminosities. However, the bursting intervals make up less than 1% of the total time, and this small duty cycle will thus reduce the average pulsed luminosity due to the bursts to a level substantially below that of the observed quasi-periodic oscillations. Thus, we conclude that the observed quasi-periodic oscillations cannot be associated with accreting pulsars in M82.

Ruling out instrumental origin. We also rule out the possibility that this signal is intrinsic to RXTE/PCA by extracting the average power density spectra of a sample of accreting supermassive black holes with PCA count rates comparable to M82. To be consistent, we used only monitoring data taken in the GoodXenon mode during the same epoch as M82. The long-term light curves of these sources in the same time range as M82 are shown in Extended Data Fig. 2. On the basis of the causality argument, active galactic nuclei with black hole masses greater than $10^6 M_{\odot}$ cannot have

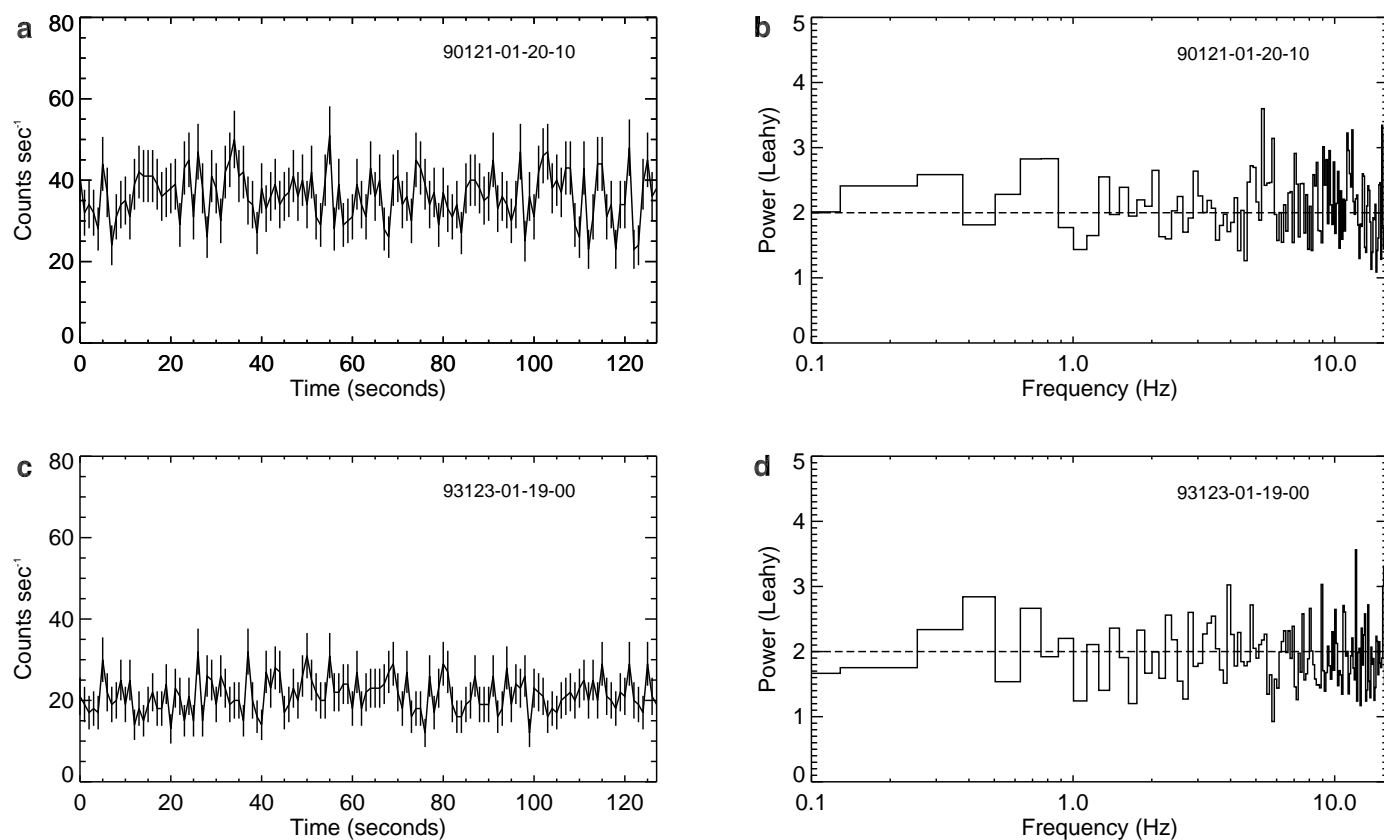
coherent oscillations at frequencies higher than ~ 0.1 Hz. The average power spectra obtained with PCA in the 3–13 keV bandpass are essentially flat and are consistent with being Poisson noise (Extended Data Fig. 3). In addition, we also extracted the average power spectrum of a blank sky field (background) monitored using the GoodXenon mode during the same epoch as M82. The corresponding average power spectrum is shown in Extended Data Fig. 4b and is again consistent with being featureless, as expected.

Relativistic precession model analysis. According to the relativistic precession model, the upper harmonic of the high-frequency quasi-periodic oscillation is associated with the Keplerian frequency at some inner radius, and the lower harmonic of the high-frequency quasi-periodic oscillation and the type-C quasi-periodic oscillation are associated with the periastron and nodal precession frequencies, respectively, at the same radius. Recently, this model was applied to GRO J1655-40²⁰, which exhibits both low-frequency and high-frequency quasi-periodic oscillations and has a very accurate mass measurement of $(5.4 \pm 0.3) M_{\odot}$ (ref. 28). They find that the black hole mass evaluated from the relativistic precession model analysis agrees nicely with its dynamical mass estimate. Given this promise of the relativistic precession model, we estimated the mass and spin of M82 X-1's black hole using this model, essentially following the methodology in ref. 20.

The relativistic precession model analysis requires that the three quasi-periodic oscillations, the two high-frequency quasi-periodic oscillations and a low-frequency quasi-periodic oscillation, be observed simultaneously. This is, however, not the case for M82 observations. Although the combined six-year RXTE/PCA data show the high-frequency quasi-periodic oscillation pair, individual XMM-Newton observations randomly dispersed over the same epoch as the RXTE monitoring have shown low-frequency quasi-periodic oscillations with frequencies in the range of 37–210 mHz (see table 2 of ref. 10 and ref. 23). Thus, we carried out the analysis for three separate quasi-periodic oscillation frequencies, the lowest and the highest values (37 and 210 mHz, respectively) and a mean quasi-periodic oscillation frequency (120 mHz).

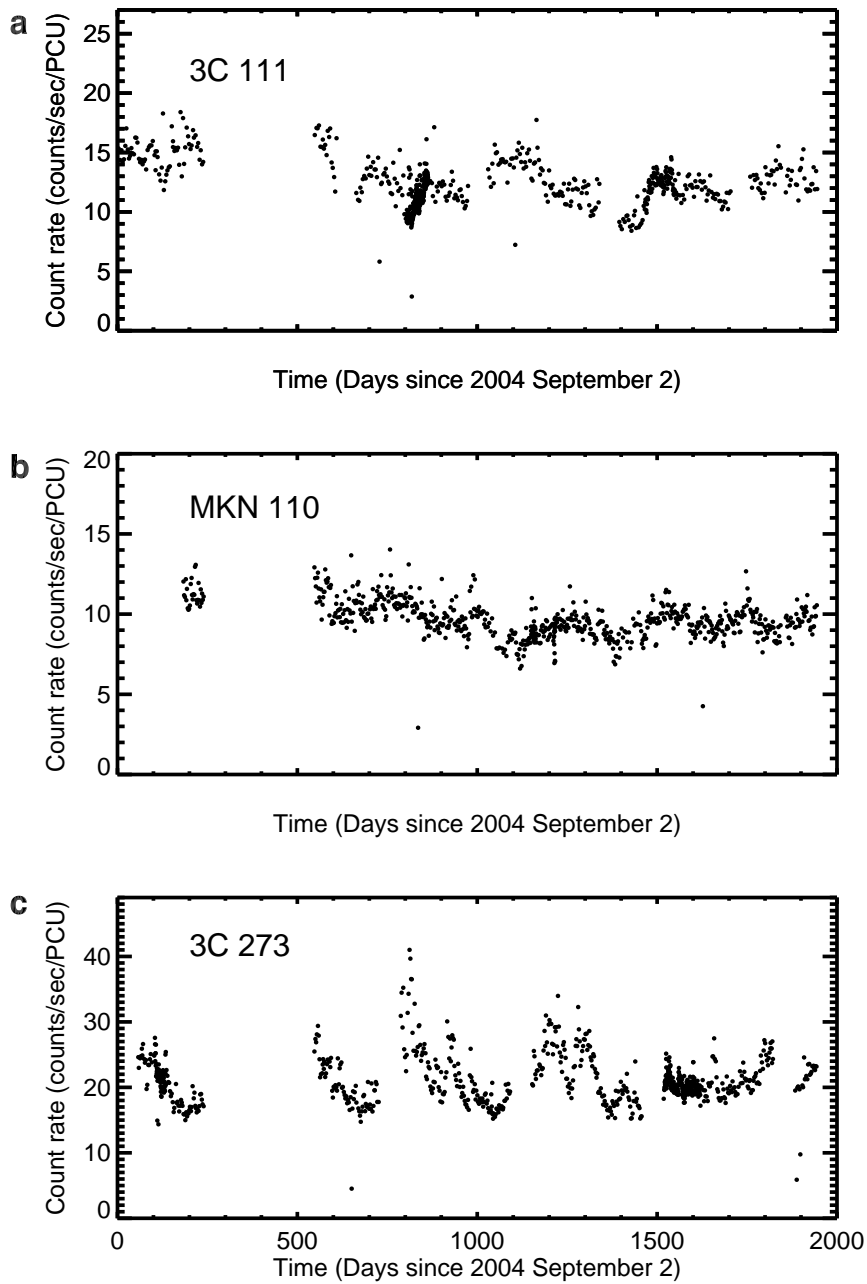
The dimensionless spin parameter is constrained to the range $0.06 < a < 0.31$, and the inferred radius, r , in the disk is in the range $5.53 < r/r_g < 6.82$ (Fig. 2).

31. van der Klis, M. in *Timing Neutron Stars: Proc. NATO Adv. Study Inst. Timing Neutron Stars* (eds Ogelman, H. & van den Heuvel, E. P. J.) 27 (Kluwer Academic/Plenum, 1989).
32. Leahy, D. A. *et al.* On searches for pulsed emission with application to four globular cluster X-ray sources - NGC 1851, 6441, 6624, and 6712. *Astrophys. J.* **266**, 160–170 (1983).
33. Chiang, Y.-K. & Kong, A. K. H. The long-term variability of the X-ray sources in M82. *Mon. Not. R. Astron. Soc.* **414**, 1329–1338 (2011).
34. Wood, K. S. *et al.* USA experiment and RXTE observations of a variable low-frequency quasi-periodic oscillation in XTE J1118+480. *Astrophys. J.* **544**, L45–L48 (2000).
35. Rodríguez, J., Corbel, S., Kalemci, E., Tomsick, J. A. & Tagger, M. An X-ray timing study of XTE J1550-564: evolution of the low-frequency quasi-periodic oscillations for the complete 2000 outburst. *Astrophys. J.* **612**, 1018–1025 (2004).
36. Jin, J., Feng, H. & Kaaret, P. Transition to the disk dominant state of a new ultraluminous X-ray source in M82. *Astrophys. J.* **716**, 181–186 (2010).
37. Feng, H., Rao, F. & Kaaret, P. Discovery of millihertz X-ray oscillations in a transient ultraluminous X-ray source in M82. *Astrophys. J.* **710**, L137–L141 (2010).
38. Laycock, S. *et al.* Long-term behavior of X-ray pulsars in the Small Magellanic Cloud. *Astrophys. J.* **161**, 96–117 (2005).
39. Giles, A. B. *et al.* The main characteristics of GRO J1744-28 observed by the proportional counter array experiment on the Rossi X-Ray Timing Explorer. *Astrophys. J.* **469**, L25 (1996).
40. Jahoda, K. *et al.* Peak luminosities of bursts from GRO J1744-28 measured with the RXTE PCA. *Nucl. Phys. B Proc. Suppl.* **69**, 210–215 (1999).
41. Sazonov, S. Y., Sunyaev, R. A. & Lund, N. Super-Eddington X-ray luminosity of the bursting pulsar GRO J1744-28: WATCH/Granat observations. *Astron. Lett.* **23**, 286–292 (1997).
42. Negoro, H. *et al.* MAXI/GSC and Swift/BAT detection of enhanced hard X-ray emission from the Galactic center region, renewed activity of GRO J1744-28? *Astron. Teleg.* **5790** (2014).
43. Finger, M. H., Jenke, P. A. & Wilson-Hodge, C. Fermi/GBM detection of 0.467s pulsation from GRO J1744-28. *Astron. Teleg.* **5810** (2014).
44. Kennea, J. A., Kouveliotou, C. & Younes, G. GRO J1744-28: Swift XRT confirmation of outburst. *Astron. Teleg.* **5845** (2014).
45. D'Ai, A. *et al.* GRO J1744-28 active as X-ray pulsar. *Astron. Teleg.* **5858** (2014).
46. Linares, M., Kennea, J., Krimm, H. & Kouveliotou, C. Swift detects bursting activity from GRO J1744-28. *Astron. Teleg.* **5883** (2014).
47. Pintore, F. *et al.* Detection of a spin derivative in GRO J1744-28 with Swift/XRT. *Astron. Teleg.* **5901** (2014).
48. Kommers, J. M. *et al.* Postburst quasi-periodic oscillations from GRO J1744-28 and from the rapid burster. *Astrophys. J.* **482**, L53 (1997).



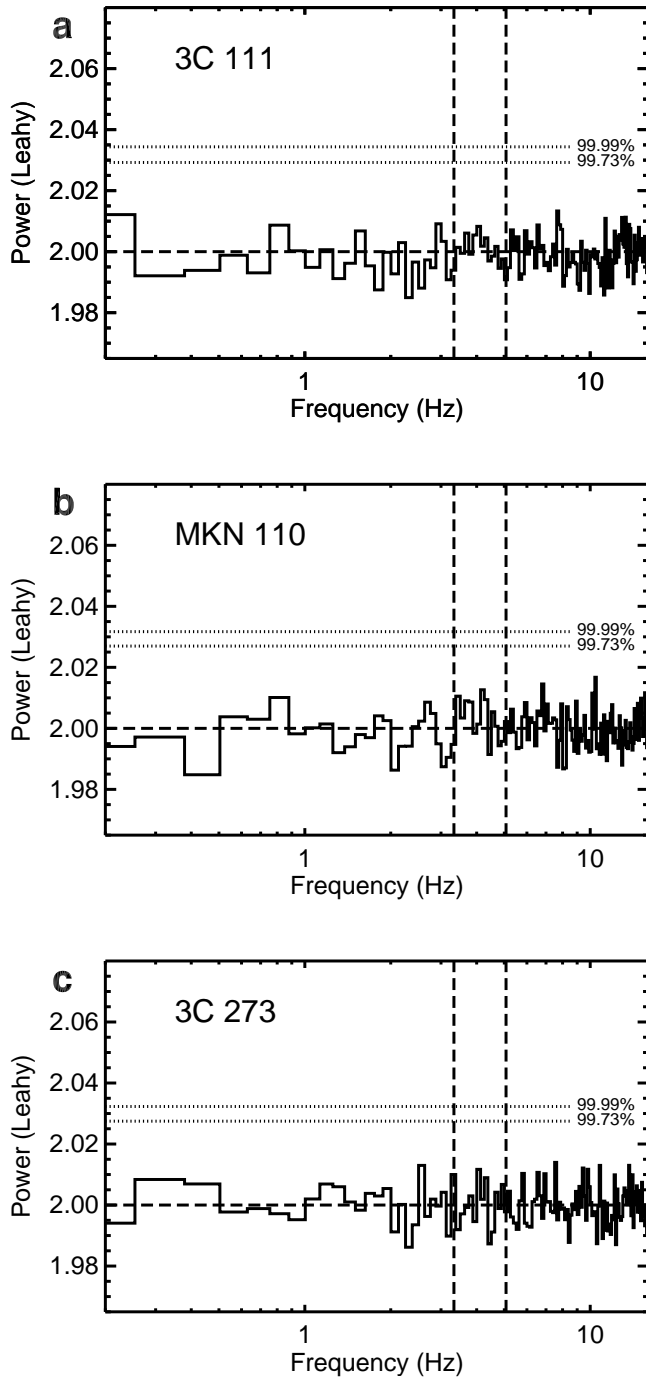
Extended Data Figure 1 | Sample RXTE/PCA light curves and power density spectra of M82. The 128 s X-ray (3–13 keV) light curves (a, c) and the corresponding power spectra (b, d) of M82. The corresponding observation

IDs are shown in the top right of each panel. The light curves have a bin size of 1 s, and the power spectra have a frequency resolution of 0.125 Hz. The error bars in a and c represent the standard error of the mean.

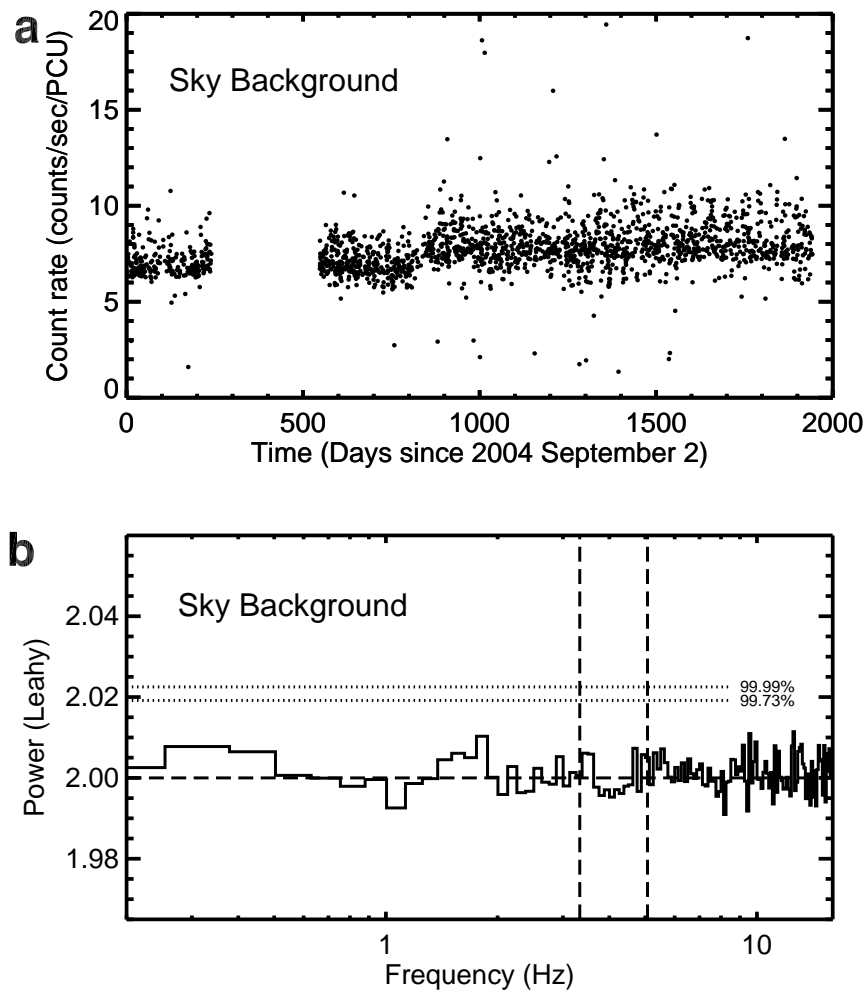


Extended Data Figure 2 | Long-term X-ray (3–13 keV) light curves of three accreting supermassive black holes. These were extracted from the same time window as the M82 observations (2004 September 2 to 2005 April 30 and

2006 March 3 to 2009 December 30). Data from 3C 111, MKN 110 and 3C 273 are shown in **a**, **b** and **c**, respectively. The count rates are not corrected for background.



Extended Data Figure 3 | Average X-ray (3–13 keV) power spectra of three accreting supermassive black holes. Similar to the M82 analysis, these spectra were extracted by combining all the data (128 s data segments) shown in Extended Data Fig. 2. The Poisson noise level is equal to 2, and the 99.73% and 99.99% confidence contours are indicated by horizontal dotted lines. The two dashed vertical lines are drawn at 3.32 and 5.07 Hz. In all three cases, 3C 111 (a), MKN 110 (b) and 3C 273 (c), there are no significant power spectral features. The total PCA exposures used for these spectra were 630 ks (a), 738 ks (b) and 711 ks (c).



Extended Data Figure 4 | Long-term X-ray light curve and the average power density spectrum of a background sky field of RXTE/PCA. Similar to Extended Data Fig. 3, the 99.73% and 99.99% contours and the vertical lines at 3.32 and 5.07 Hz are indicated in the power density spectrum (b). Again, only

observations coincident with M82 monitoring were used (2004 September 2 to 2005 April 30 and 2006 March 3 to 2009 December 30), and the total exposure time was 1,450 ks. The light curve (3–13 keV) is shown in (a). The background has coordinates of $RA = 5.0^\circ$, $dec. = -67.0^\circ$.

Extended Data Table 1 | Dependence of the percentage r.m.s. amplitudes of the two oscillations in the X-ray bandpass

PCA Channels	Energy Range	Net Count Rate ^a	Background Count Rate	Uncorrected % rms Amplitude ^b	Corrected % rms Amplitude ^c
3.32 Hz Quasi-Periodic Oscillation					
7-18	3-8 keV	19.0	10.5	1.1±0.2	2.8±0.4
7-24	3-10 keV	24.3	14.5	1.1±0.2	2.5±0.4
7-32	3-13 keV	29.9	18.9	1.1±0.1	2.7±0.4
5.07 Hz Quasi-Periodic Oscillation					
7-18	3-8 keV	19.0	10.5	1.2±0.2	2.5±0.4
7-24	3-10 keV	24.3	14.5	1.0±0.2	2.8±0.4
7-32	3-13 keV	29.9	18.9	1.0±0.1	3.0±0.4

* The total (source plus background) count rate in the given energy range (a).

† Not corrected for the background (b).

‡ Background-corrected percentage r.m.s. amplitude (c) where corrected r.m.s. amplitude = (uncorrected r.m.s. amplitude) × (total count rate)/(source count rate). The source count rate is the total rate minus the background rate. We used 128 s data segments for this study.

Hierarchical organization of chiral rafts in colloidal membranes

Prerna Sharma^{1*}, Andrew Ward^{1*}, T. Gibaud², Michael F. Hagan¹ & Zvonimir Dogic¹

Liquid–liquid phase separation is ubiquitous in suspensions of nanoparticles, proteins and colloids. It has an important role in gel formation, protein crystallization and perhaps even as an organizing principle in cellular biology^{1,2}. With a few notable exceptions^{3,4}, liquid–liquid phase separation in bulk proceeds through the continuous coalescence of droplets until the system undergoes complete phase separation. But when colloids, nanoparticles or proteins are confined to interfaces, surfaces or membranes, their interactions differ fundamentally from those mediated by isotropic solvents^{5,6}, and this results in significantly more complex phase behaviour^{7–13}. Here we show that liquid–liquid phase separation in monolayer membranes composed of two dissimilar chiral colloidal rods gives rise to thermodynamically stable rafts that constantly exchange monomeric rods with the background reservoir to maintain a self-limited size. We visualize and manipulate rafts to quantify their assembly kinetics and to show that membrane distortions arising from the rods' chirality lead to long-range repulsive raft–raft interactions. Rafts assemble into cluster crystals at high densities, but they can also form bonds to yield higher-order structures. Taken together, our observations demonstrate a robust membrane-based pathway for the assembly of monodisperse membrane clusters that is complementary to existing methods for colloid assembly in bulk suspensions^{14–16}. They also reveal that chiral inclusions in membranes can acquire long-range repulsive interactions, which might more generally have a role in stabilizing assemblages of finite size^{13,17}.

In the presence of a non-adsorbing polymer, monodisperse rod-like viruses experience effective depletion attractions that drive their lateral

association. These interactions can lead to the assembly of colloidal monolayer membranes, one rod-length thick, that are held together by the osmotic pressure of the enveloping polymer suspension¹⁸. Colloidal monolayers and lipid bilayers exhibit long-wavelength fluctuations described by the same free energy, despite their molecular-scale differences.

In this work we investigated the behaviour of colloidal membranes containing a mixture of two rods: rod-like *fd*-Y21M virus (880 nm long) and M13KO7 virus (1,200 nm long)¹⁹. Membranes were assembled by adding a depletant to a dilute isotropic mixture of *fd*-Y21M and M13KO7. After reaching a large enough size, membranes sedimented to the bottom of the sample chambers; the constituent rods pointed in the *z* direction and images were taken in the *x*–*y* plane (Fig. 1). For all parameters investigated, both rods co-assembled into the same colloidal membrane. At low depletant concentration the rods remained homogeneously mixed throughout the membrane (Fig. 1b, d), whereas at high depletant concentration they separated into bulk phases (Fig. 1b, f). At intermediate concentrations we observed the formation of highly monodisperse micrometre-sized droplets (colloidal rafts) enriched in short rods and floating in the background of long rods (Fig. 1e). Colloidal rafts did not coarsen with time, suggesting that they were equilibrium structures.

To test this hypothesis we brought two membranes, each containing only one type of virus, into close proximity (Fig. 1a). On coalescence, the interface separating dissimilar rods became unstable as short rods invaginated into the long-rod fluid, giving rise to an interconnected network of rafts held together by thin metastable bridges (Fig. 1c and Supplementary Video 1). Over time the conjoining liquid bridges disintegrated, leading to the formation of isolated finite-sized rafts and demonstrating

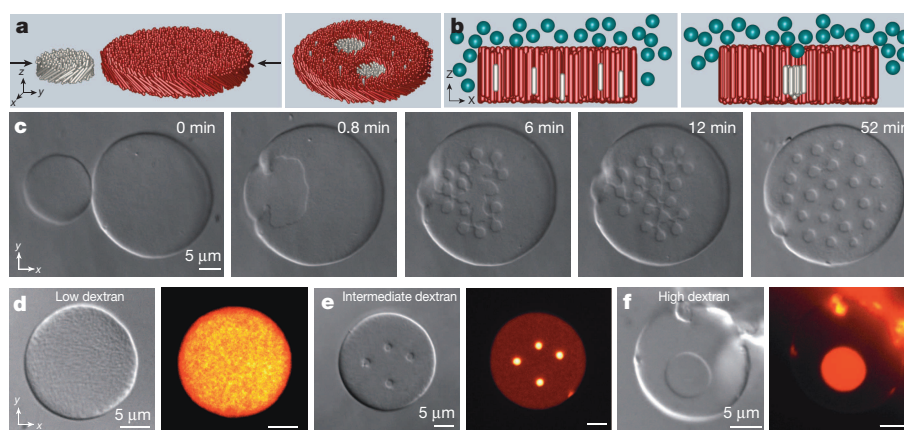


Figure 1 | Membrane-mediated assembly of monodisperse chiral colloidal rafts. **a**, Coalescence of membranes composed of short and long rods of opposite chiralities. **b**, Clustering of short rods dissolved in a long-rod membrane is driven by excluded-volume interactions. **c**, Coalescence of short right-handed *fd*-Y21M and long left-handed M13KO7 membranes leads to the formation of finite-sized rafts. **d**, Differential interference contrast and fluorescence images of a *fd*-Y21M–M13KO7 membrane at 34 mg ml⁻¹

dextran concentration show homogeneous mixing. *fd*-Y21M is fluorescently labelled. **e**, Formation of finite-sized clusters enriched in fluorescently labelled *fd*-Y21M at intermediate dextran concentrations (38 mg ml⁻¹). The background hue indicates that a small fraction of *fd*-Y21M rods are dissolved in the background membrane. **f**, Complete separation of the bulk phase is observed at the highest dextran concentrations (52 mg ml⁻¹).

¹Department of Physics, Brandeis University, 415 South Street, Waltham, Massachusetts 02454, USA. ²Laboratoire de Physique, École Normale Supérieure de Lyon, Université de Lyon I, CNRS/UMR 5672, 46 allée d'Italie, 69007 Lyon, France.

*These authors contributed equally to this work.

that colloidal rafts have a lower free energy than a bulk separated phase. We have explored the behaviour of six phase-separating mixtures composed of viruses of various physical properties; in each case we observed micro-separated phases at intermediate polymer concentrations, suggesting that rafts are a ubiquitous feature of colloidal membranes (Methods). Here we focus on the behaviour of the *fd*-Y21M–M13K07 mixture, using the unique features of colloidal membranes that allow us to manipulate rafts optically, to visualize their nucleation pathways and to track the real-space dynamics of isolated constituent rods.

We determined the raft free-energy landscape by measuring how k_{on} and k_{off} , the rates at which rods associate and dissociate from a raft, depend on the raft size. We measured k_{off} using fluorescence recovery after photobleaching (FRAP)²⁰. Using a focused beam, we bleached an isolated raft of predetermined size. The fluorescence signal recovered within minutes, indicating an exchange of bleached raft-bound rods with unbleached background rods (Fig. 2a and Supplementary Video 2). The fluorescence recovery curves were fitted by single-exponential curves (Fig. 2e). If the raft size does not change appreciably during of experiment, the fluorescence recovery time constant is $1/k_{\text{off}}$ (Methods). By performing FRAP measurements on rafts of varying radii we determined the dependence of k_{off} on raft size (Fig. 2h).

To measure k_{on} we created a raft population with heterogeneous radii R and quantified their subsequent growth rates (Fig. 2b, c, f and Supplementary Video 3). Rafts with $R < R_{\text{critical}}$ quickly evaporated into the background membrane (Fig. 2g). When $R_{\text{equilibrium}} > R > R_{\text{critical}}$, rafts grew until attaining the equilibrium size. Finally, rafts with $R > R_{\text{equilibrium}}$

shrank slowly until achieving equilibrium. The evolution of raft size is governed by the kinetic equation

$$\frac{dN(t)}{dt} = -k_{\text{off}}N(t) + k_{\text{on}}C_{\text{BG}}$$

where $N(t)$ is the number of raft-bound rods and C_{BG} is the concentration of short fluorescently labelled rods in the background. $dN(t)/dt$ is determined from size evolution experiments (Fig. 2g), k_{off} is determined through FRAP experiments, and C_{BG} is proportional to the mean fluorescence of the background membrane. The only unknown is k_{on} , making it possible to determine its dependence on raft size. We found that k_{on} increased weakly with increasing raft size, consistent with two-dimensional diffusion-limited association kinetics. Having measured both k_{on} and k_{off} (Fig. 2h), we calculate $\Delta G(N) = k_{\text{B}}T(k_{\text{off}}NA_{\text{rod}}/k_{\text{on}})$, where NA_{rod} is the raft area and $\Delta G(N)$ is the energy cost of adding a single particle to a raft (Fig. 2i). The minimum of the excess free energy satisfies $\Delta G - k_{\text{B}}T\ln(C_{\text{BG}}A_{\text{rod}}) = 0$ and determines the equilibrium raft size, which corresponds to $\sim 24,000$ rods. Integrating ΔG over N reveals how the raft free energy, $G(N)$, scales with the number of rods. The width of the parabola determines the fluctuations in raft area, namely $\sim 1,400$ particles, corresponding to fluctuations of 30 nm in radius. Such fluctuations are below the resolution limit of the optical microscope, consistent with our observation of a monodisperse raft distribution in equilibrium.

To validate our analysis independently, we performed single-molecule lifetime experiments by tracking a small fraction of fluorescently labelled

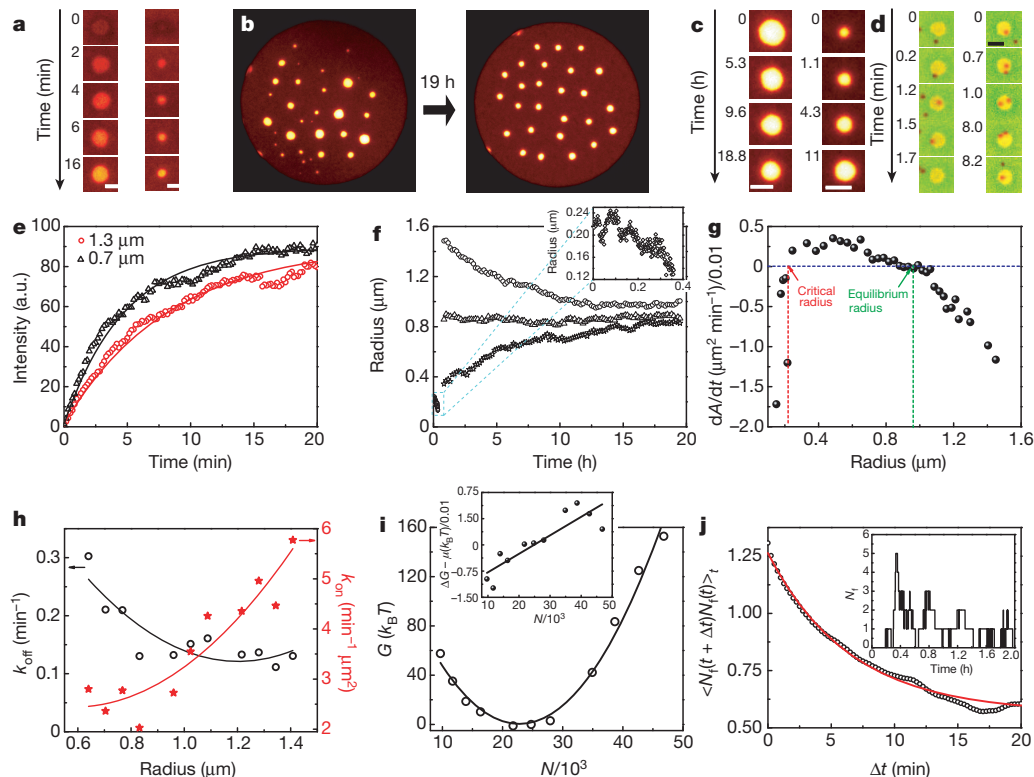


Figure 2 | Single-molecule analysis reveals raft nucleation dynamics and assembly kinetics. **a**, The fluorescence of a bleached cluster recovers after a few minutes, revealing kinetics of rod exchange between rafts and the membrane background. **b**, Initially polydisperse rafts equilibrate in size after many hours. **c**, Expansion of an undersized raft and contraction of an oversized raft. **d**, Fluorescently labelled *fd*-Y21M rods associate and dissociate from a raft, revealing binding kinetics at the single-molecule level. **e**, Exponential recovery of cluster fluorescence intensity after a photobleaching event yields k_{off} for rafts with different sizes. **f**, Time evolution of cluster growth depends on the initial cluster size. Inset: a subcritical raft quickly evaporates into the membrane background. **g**, The dependence of raft expansion rates on raft size directly

reveals the critical nucleus size and equilibrium size. **h**, Combining single-raft fluorescence recovery measurements with raft expansion rates reveals the dependence of k_{off} (circles) and k_{on} (stars) on raft size (each point represents a single measurement). **i**, The minimum in raft free-energy landscape indicates the equilibrium raft size. Inset: ΔG , the free-energy change associated with adding a single virus to a colloidal raft (each point represents a single measurement). **j**, Exponential decay of the autocorrelation function of the number of raft-bound rods, $N_i(t)$, yields both k_{off} and k_{on} . For equilibrium-sized rafts, single-molecule analysis agrees with measurements in **h**. Inset: plot of $N_i(t)$. Scale bars, 2.5 μm .

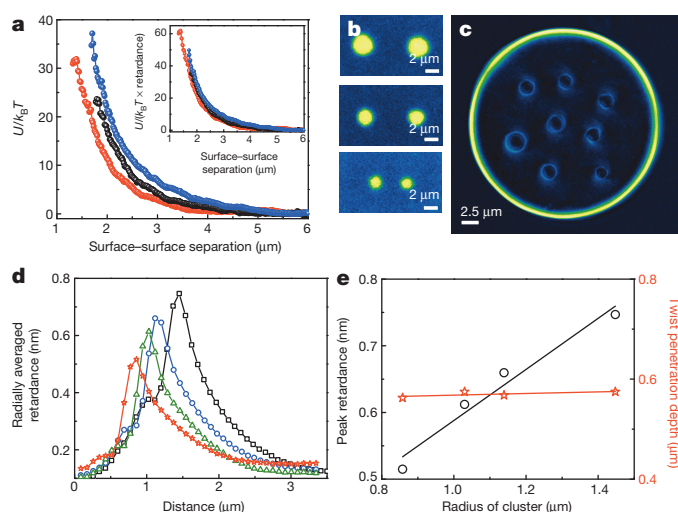


Figure 3 | Raft-induced chiral twist governs membrane-mediated raft repulsions. **a**, Effective pair interaction potential obtained using the blinking trap technique for clusters with diameters of 2.9 μm (blue), 2.1 μm (black) and 1.6 μm (red). Inset: interaction potential rescaled by the maximum tilt angle. **b**, Images of cluster pairs for which interactions were measured. **c**, LC-PolScope image of heterogeneous rafts embedded in a membrane. **d**, Dependence of radially averaged retardance profile on raft size. **e**, Peak retardance (maximum tilt angle; circles) increases linearly with increasing raft size. The twist penetration depth (stars) is independent of domain size (each point represents a single measurement).

short rods in an otherwise unlabelled membrane (Fig. 2d and Supplementary Video 4). Overlaying of fluorescent images onto simultaneously acquired phase-contrast images revealed raft-bound rods as a function of time, $N_r(t)$ (Fig. 2j). From such results we find that for equilibrium rafts $k_{\text{on}} = 3.7 \text{ min}^{-1} \mu\text{m}^2$ and $k_{\text{off}} = 0.15 \text{ min}^{-1}$, which is in agreement with FRAP measurements (Methods). Very few other experimental systems^{14,21} permit the visualization of subcritical nucleation dynamics and the measurement of the associated free-energy landscape enabled by our colloidal membranes.

With an understanding of single-raft assembly kinetics, we subsequently measured membrane-mediated raft–raft interactions by adapting a technique based on blinking optical tweezers²². To manipulate rafts, which are repelled from a focused laser, we created an optical plough composed of multiple beams. Using two ploughs we pushed a raft pair into close proximity. Subsequently, we switched off the traps and quantified raft trajectories as they were pushed apart by repulsive interactions (Supplementary Video 5). Repeating this procedure dozens of times yielded the inter-raft potential (Methods). The effective repulsive potential is described by an exponential with a characteristic length scale of 0.65 μm (Fig. 3a). The measured raft–raft interactions can be quantitatively related to raft-induced membrane distortions visualized with a quantitative polarized light microscope (LC-PolScope). This technique yields images whose intensity indicates the local retardance, which is related to the local tilt of rods away from the membrane normal vector (Fig. 3c)²³. The interior of a homogeneous one-component membrane viewed in the x – y plane lacks optical anisotropy and appears dark in the LC-PolScope. In contrast, rafts imaged with the LC-PolScope had a spatially varying retardance, indicating local twisting of raft-bound rods.

Such deformations can be understood by considering that *fd*-Y21M and M13KO7 have right-handed and left-handed chirality, respectively^{19,24}. As a consequence, the interaction energy between membrane-embedded viruses is minimized when neighbouring rods are not parallel to each other but are twisted at a small but finite angle. The radially averaged retardance profile reveals that, starting from the raft centre, *fd*-Y21M rods twist with a right-handed sense, leading to increasing rod tilt away from the monolayer normal. The tilt attains a maximum at the raft's edge.

Moving past the edge, the membrane is enriched in left-handed M13KO7. The left-handed twist induces tilting of rods back towards the membrane normal. Larger rafts attain a larger maximum edge tilt; however, the characteristic length over which twist penetrates into the background membrane is independent of raft size (Fig. 3d, e). The length scales associated with raft repulsion and raft-induced twist deformation are comparable, suggesting a possible link.

Homo-chiral colloidal membranes are inherently frustrated, because the constituent rods cannot simultaneously twist locally and assemble into a monolayer globally. Consequently, twist is expelled from the membrane interior to the edges. Increasing the rod chirality raises the free energy of interior untwisted rods while lowering the free energy of edge-bound twisted rods, leading to chiral control of edge line tension^{25,26}. For uniformly mixed binary membranes, the same constraints enforce untwisting of all constituent chiral rods, thus raising their free energy. However, the formation of right-handed rafts in the left-handed background allows twist to penetrate into the membrane interior, lowering raft free energy. Besides explaining raft stability, this picture also explains the origin of inter-raft repulsion. Bringing two rafts together closer than the twist penetration length requires energetically costly untwisting of chiral rods. Our hypothesis predicts that the range of raft repulsion is independent of cluster size, while its strength increases linearly with cluster size. Measured potentials between clusters of different sizes agree with these predictions (Fig. 3a, b). Furthermore, when the magnitude of the repulsive interactions is rescaled by the maximum twist at the rafts edge obtained from LC-PolScope images, all potentials collapse on top of each other, thus establishing a quantitative relationship between raft-induced chiral distortions and membrane-mediated raft interactions. These experiments describe how chirality stabilizes finite-sized rafts, a mechanism that is complementary to other proposed mechanisms^{27–29}.

The emergent raft repulsions can be used to assemble higher-order structures. To explore this capability, we measured the raft radial distribution function, $g(r)$, for samples with increasing raft density (Supplementary Video 6). In the dilute limit, $g(r)$ is zero for small separations, indicative of strong repulsive interactions, and constant elsewhere (Fig. 4a). At intermediate raft densities, $g(r)$ exhibits liquid-like oscillations that decay after a few coordination shells (Fig. 4b). At highest densities, rafts form two-dimensional crystal-like structures as demonstrated by sharp peaks in $g(r)$ (Fig. 4c). This behaviour is reminiscent of conventional two-dimensional repulsive colloids. However, there are also important differences, because colloidal crystals are assembled from immutable solid particles whereas the membrane-embedded crystals are assembled from highly adaptable, size-adjustable rafts, each of which constantly exchanges rods with the background. Moreover, whereas conventional droplets in bulk suspensions always assume a surface-minimizing spherical shape, the behaviour of rafts embedded in a membrane is significantly more complex. For example, at higher ionic strength (125 mM) and depletant concentration (40 mg ml^{-1} and above), the interactions between rafts are no longer purely repulsive. Instead, as two rafts approach each other they can form a stable bond that takes a shape of a thin liquid bridge. Such processes lead to the formation of raft dimers, trimers and higher-order structures, which remain stable for many days (Fig. 4d and Supplementary Video 7). It is even possible to observe the assembly of highly complex structures such as a classical bead-spring polymer formed from nine raft subunits. The resulting polymer-like structure exhibits large fluctuations in shape; a quantitative analysis of these fluctuations reveals that the preferred angle of the neighbouring bonds is 109° (Fig. 4e and Supplementary Video 8).

Thus, we have demonstrated that chirality is important in the formation of thermodynamically stable, finite-sized, monodisperse rafts and their assembly into higher-order structures through membrane-mediated repulsive interactions that are orders of magnitude longer in range than attractive depletion interactions³⁰. The chiral molecule cholesterol is known to be essential for the assembly of lipid rafts in biological membranes. However, our membranes comprised two oppositely twisted chiral colloids, whereas biological membranes contain only one

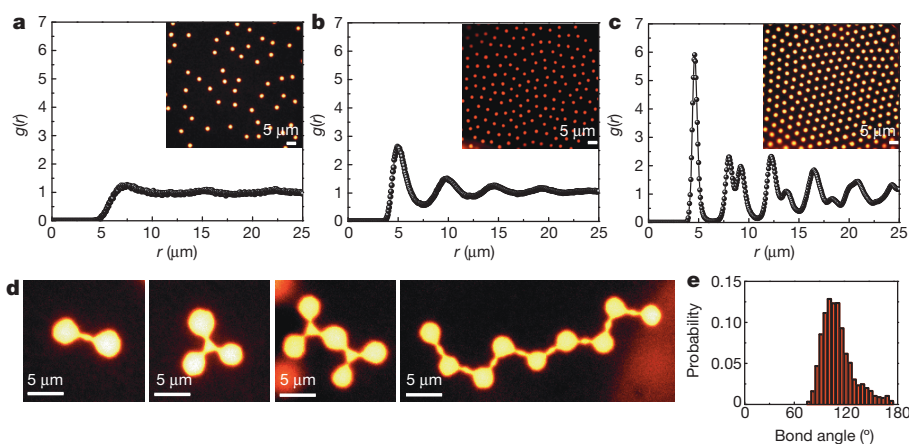


Figure 4 | Assembly of raft crystals and aspherical supra-rafts. **a**, Radial distribution function, $g(r)$, of dilute membrane-embedded clusters. The stoichiometric ratio of M13KO7 to *fd*-Y21M is 6:1. Inset: representative image of the membrane. **b**, At intermediate raft densities $g(r)$ exhibits a liquid-like structure (ratio of M13KO7 to *fd*-Y21M 4:1). **c**, At high densities, rafts organize into a two-dimensional cluster crystal (ratio of M13KO7 to *fd*-Y21M 2.2:1).

chiral constituent; the possible biological relevance of the chiral effects we have documented therefore remains an open question.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 May; accepted 14 July 2014.

- ten Wolde, P. R. & Frenkel, D. Enhancement of protein crystal nucleation by critical density fluctuations. *Science* **277**, 1975–1978 (1997).
- Hyman, A. A. & Simons, K. Beyond oil and water-phase transitions in cells. *Science* **337**, 1047–1049 (2012).
- Stradner, A. *et al.* Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature* **432**, 492–495 (2004).
- Groenewold, J. & Kegels, W. K. Anomalous large equilibrium clusters of colloids. *J. Phys. Chem. B* **105**, 11702–11709 (2001).
- Goulian, M., Bruinsma, R. & Pincus, P. Long-range forces in heterogeneous fluid membranes. *Europhys. Lett.* **23**, 155 (1993).
- Dan, N., Berman, A., Pincus, P. & Safran, S. A. Membrane-induced interactions between inclusions. *J. Phys. II* **4**, 1713–1725 (1994).
- Weis, R. M. & McConnell, H. M. Two-dimensional chiral crystals of phospholipid. *Nature* **310**, 47–49 (1984).
- Dietrich, C. *et al.* Lipid rafts reconstituted in model membranes. *Biophys. J.* **80**, 1417–1428 (2001).
- Dinsmore, A. D. *et al.* Colloidosomes: selectively permeable capsules composed of colloidal particles. *Science* **298**, 1006–1009 (2002).
- Lin, Y., Skaff, H., Emrick, T., Dinsmore, A. D. & Russell, T. P. Nanoparticle assembly and transport at liquid–liquid interfaces. *Science* **299**, 226–229 (2003).
- Veatch, S. L. & Keller, S. L. Separation of liquid phases in giant vesicles of ternary mixtures of phospholipids and cholesterol. *Biophys. J.* **85**, 3074–3083 (2003).
- Baumgart, T., Hess, S. T. & Webb, W. W. Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Nature* **425**, 821–824 (2003).
- Lingwood, D. & Simons, K. Lipid rafts as a membrane-organizing principle. *Science* **327**, 46–50 (2010).
- Meng, G. N., Arkus, N., Brenner, M. P. & Manoharan, V. N. The free-energy landscape of clusters of attractive hard spheres. *Science* **327**, 560–563 (2010).
- Chen, Q. *et al.* Supracolloidal reaction kinetics of Janus spheres. *Science* **331**, 199–202 (2011).
- Wang, Y. F. *et al.* Colloids with valence and specific directional bonding. *Nature* **491**, 51–55 (2012).
- Sarasij, R. C., Mayor, S. & Rao, M. Chirality-induced budding: a raft-mediated mechanism for endocytosis and morphology of caveolae? *Biophys. J.* **92**, 3140–3158 (2007).
- Barry, E. & Dogic, Z. Entropy driven self-assembly of nonamphiphilic colloidal membranes. *Proc. Natl Acad. Sci. USA* **107**, 10348–10353 (2010).

d, Rafts can be permanently bonded by thin liquid bridges to form complex architectures. Fluorescence images of a raft dimer, trimer, pentamer and polymer-like nonamer are shown. **e**, Probability distribution of bond angles in the bead-spring nonamer obtained by analysing 400 independent configurations of the fluctuating polymer.

- Barry, E., Beller, D. & Dogic, Z. A model liquid crystalline system based on rodlike viruses with variable chirality and persistence length. *Soft Matter* **5**, 2563–2570 (2009).
- Sprague, B. L., Pego, R. L., Stavreva, D. A. & McNally, J. G. Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophys. J.* **86**, 3473–3495 (2004).
- Gasser, U., Weeks, E. R., Schofield, A., Pusey, P. N. & Weitz, D. A. Real-space imaging of nucleation and growth in colloidal crystallization. *Science* **292**, 258–262 (2001).
- Crocker, J. C. & Grier, D. G. Microscopic measurement of the pair interaction potential of charge-stabilized colloid. *Phys. Rev. Lett.* **73**, 352–355 (1994).
- Barry, E., Dogic, Z., Meyer, R. B., Pelcovits, R. A. & Oldenbourg, R. Direct measurement of the twist penetration length in a single smectic A layer of colloidal virus particles. *J. Phys. Chem. B* **113**, 3910–3913 (2009).
- Tombolato, F., Ferrarini, A. & Grelet, E. Chiral nematic phase of suspensions of rodlike viruses: left-handed phase helicity from a right-handed molecular helix. *Phys. Rev. Lett.* **96**, 101103/PhysRevLett.96.258302 (2006).
- Gibaud, T. *et al.* Reconfigurable self-assembly through chiral control of interfacial tension. *Nature* **481**, 348–351 (2012).
- Kaplan, C. N. & Meyer, R. B. Colloidal membranes of hard rods: unified theory of free edge structure and twist walls. *Soft Matter* **10**, 4700–4710 (2014).
- Rozovsky, S., Kaizuka, Y. & Groves, J. T. Formation and spatio-temporal evolution of periodic structures in lipid bilayers. *J. Am. Chem. Soc.* **127**, 36–37 (2005).
- Reynwar, B. J. *et al.* Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature* **447**, 461–464 (2007).
- Ursell, T. S., Klug, W. S. & Phillips, R. Morphology and interaction between lipid domains. *Proc. Natl Acad. Sci. USA* **106**, 13301–13306 (2009).
- Seul, M. & Andelman, D. Domain shapes and patterns—the phenomenology of modulated phases. *Science* **267**, 476–483 (1995).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge discussions with R. B. Meyer. This work was supported by the US National Science Foundation (NSF-MRSEC-0820492, NSF-DMR-0955776 and NSF-MRI-0923057) and the Petroleum Research Fund (ACS-PRF 50558-DNI7). We acknowledge use of the Brandeis MRSEC optical microscopy facility.

Author Contributions P.S., A.W. and Z.D. conceived and designed the experiments. P.S. and A.W. performed the experiments. M.F.H. designed the theoretical models. P.S., A.W., Z.D. and M.F.H. interpreted the experiments. T.G. provided material samples. P.S. and Z.D. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Z.D. (zdogic@brandeis.edu).

METHODS

Virus growth and sample preparation. For our experiments we prepared a number of different viruses of varying contour length, chirality and persistence length including: *fd*-wt (wild type), *fd*-Y21M, M13KO7, M13-wt (wild type), M13mp18, *fd*-tet and Litmus phagemid grown with M13KO7 helper phage¹⁹. In preliminary investigations we explored six different mixtures: *fd*-Y21M and M13KO7, M13-wt and M13KO7, M13mp18 and M13KO7, *fd*-wt and *fd*-tet, *fd*-wt and M13KO7, and *fd*-wt and Litmus phagemid. In all these mixtures we found that stable microphase separation was located between uniformly mixed membranes at low depletant concentrations and bulk separation at high concentration. To make progress we primarily focused on a mixture of filamentous bacteriophages *fd*-Y21M and M13KO7. The rod-like *fd*-Y21M virus has a diameter of 6 nm, a contour length of 880 nm and a persistence length of 9.9 μm (ref. 19). M13KO7 has the same diameter but a contour length of 1,200 nm and a persistence length of 2.8 μm (ref. 31). Both systems exhibit isotropic, cholesteric and smectic phases with increasing virus concentration. However, *fd*-Y21M forms a cholesteric phase with a right-handed twist, whereas the M13KO7 cholesteric phase twists in the opposite direction^{19,24}.

M13KO7 and *fd*-Y21M viruses were purified using standard biological protocols³². To remove end-to-end dimers or longer viruses that destabilize colloidal membranes, purified virus suspensions were fractionated through the isotropic–nematic phase transition²⁰. Only the isotropic fraction, enriched in nominal length viruses, was used for assembling colloidal membranes. Viruses were suspended in Tris-HCl buffer (20 mM pH 8.0) to which 100 mM NaCl was added to screen electrostatic interactions. For fluorescence visualization, primary amines of the major coat protein of *fd*-Y21M were labelled with amine reactive fluorophore (DyLight-NHS ester 550; Thermo Fisher)³³. There are about 2,700 labelling sites available on the virus surface. However, each virus was labelled at a low volume fraction (~ 30 dye molecules per virus), ensuring that the labelling did not affect the system phase behaviour. *fd*-Y21M and M13KO7 were mixed at known stoichiometric ratios. The non-adsorbing polymer dextran (molecular mass 500,000 Da; Sigma-Aldrich) was added to this suspension, and the resultant suspension was injected into an observation chamber. The chamber was made from a microscope slide and a coverslip (Goldseal; Fisher Scientific) separated by a spacer consisting of a non-stretched Parafilm and sealed with ultraviolet-cured glue (Norland Optical). Glass surfaces were thoroughly cleaned with a hot 1% soap solution (Hellmanex; Hellma Analytics). Samples prepared in this manner remained stable for weeks or even months. The final concentration of bi-disperse virus mixture was 2.5 mg ml^{-1} . All the measurements were performed for membranes assembled at 40 mg ml^{-1} dextran concentration. Membranes initially assembled in the bulk suspension but owing to their density they eventually sedimented onto the coverslip. To prevent non-specific absorption, glass surfaces were coated with a polyacrylamide brush, which suppressed the depletion interaction between viruses and the glass walls³⁴.

Optical microscopy methods. Within a liquid-like colloidal monolayer membrane, rods with opposite chirality phases separated to form finite-sized colloidal rafts, with the exact behaviour depending on the depletant concentration. Raft structures and dynamics were examined with several complementary optical microscopy techniques. Differential interference contrast, phase contrast and fluorescence images were acquired using an inverted microscope (Nikon TE2000) equipped with an oil-immersion objective (1.3 numerical aperture, $\times 100$ Plan-Fluor) and connected to charge-coupled-device camera (Andor Clara). Fluorescently labelled *fd*-Y21M virus was imaged using the same setup equipped with mercury–halide epifluorescence source and a rhodamine filter cube (excitation wavelength 532–554 nm, emission wavelength 570–613 nm). The exposure time for acquisition of fluorescent images was 100 ms.

Spatial variations of rod tilt with respect to the membrane normal were determined using an LC-PolScope (Cambridge Research and Instrumentation)³⁵. The LC-PolScope yields images in which the intensity of each pixel is proportional to local sample birefringence. Subtraction of the background birefringence enabled the visualization of subnanometre birefringent images with LC-PolScope, something that is not possible with conventional techniques. For a membrane lying in the image plane, it is possible to translate the LC-PolScope birefringence signal quantitatively into local tilting of the rods away from the membrane normal^{23,36}.

Determination of k_{off} using single-raft FRAP. DyLight 550 attached to *fd*-Y21M surface has an excitation wavelength of 562 nm. An isolated raft of predetermined size enriched with fluorescently labelled *fd*-Y21M was illuminated with a focused laser beam (~ 100 mW, 514 nm). Under these conditions, raft-bound rods bleached instantaneously for all practical purposes. The exchange of rods between rafts and the background membrane caused fluorescence to recover on a timescale of minutes. This process can be modelled using kinetic equations. $N_f(t)$ and $N_b(t)$, respectively, describe the number of fluorescent and bleached rods in a raft at time t . Assuming that the raft size does not change on fluorescence recovery timescales and that the concentration of bleached rods in the membrane background is negligible, the time evolution of $N_f(t)$ and $N_b(t)$ is given by

$$\frac{dN_f(t)}{dt} = -k_{\text{off}}N_f(t) + k_{\text{on}}C_{\text{BG}}$$

$$\frac{dN_b(t)}{dt} = -k_{\text{off}}N_b(t)$$

where k_{off} and k_{on} are rod dissociation and association rates and C_{BG} is the concentration of fluorescent *fd*-Y21M rods in the background membrane²⁰. The experimentally measured total raft intensity is given by $I_{\text{tot}}(t) = \alpha N_f(t) + \beta N_b(t)$, where α and β are the intensities of single fluorescently labelled and bleached viruses, respectively. Perfect bleaching would result in $\beta = 0$. However, we found that this not the case: bleached rods retained finite fluorescence. Solving for $I_{\text{tot}}(t)$ yields

$$I_{\text{tot}}(t) = \exp(-k_{\text{off}}t) \left(I_0 - I_{\text{av}}^{\text{BG}} k_{\text{on}}/k_{\text{off}} \right) + I_{\text{av}}^{\text{BG}} k_{\text{on}}/k_{\text{off}} \quad (1)$$

where I_0 is the raft fluorescence intensity after completion of the photo-bleaching event and $I_{\text{av}}^{\text{BG}}$ is the mean fluorescence intensity of the background membrane. Equation (1) shows that the fluorescence recovery process can be approximated by a single-exponential rate, in agreement with experimental findings. The exponential time constant yields the rod dissociation constant $1/k_{\text{off}}$ (Fig. 2e, h).

Evolution of polydisperse rafts yields k_{on} . Heterogeneously sized rafts evolved towards equilibrium, a state characterized by rafts of uniform size. Using optical tweezers we fragmented/fused a small fraction of rafts, thus creating an artificially polydisperse population, and then quantified their subsequent size evolution. The long-time-lapse videos required for this measurement were acquired on a microscope equipped with auto-focus capabilities (Perfect-Focus; Nikon Ti-E). The rate equation governing the evolution of raft size is

$$\frac{dN(t)}{dt} = -k_{\text{off}}N(t) + k_{\text{on}}C_{\text{BG}} \quad (2)$$

where $N(t)$ is the total number of raft bound rods at time t . Cluster size evolution experiments yielded $dN(t)/dt$ and $N(t)$ (Fig. 2g). C_{BG} was extracted from fluorescent membrane images, and k_{off} and its size were obtained from previously described single-raft FRAP measurements. It follows that k_{on} and its dependence on raft size can be uniquely determined from equation (2), because all other parameters are determined from independent experiments.

Measuring the raft free-energy landscape. The change in free energy, ΔG , required to add a single rod to a raft containing N rods is given by

$$k_{\text{on}}/k_{\text{off}} = A_{\text{rod}}N e^{-\Delta G/k_B T} \quad (3)$$

where A_{rod} is the effective area occupied by a single *fd*-Y21M virus, N is the number of rods in a raft and $A_{\text{rod}}N$ is the raft size. Assuming dilute suspension conditions, the chemical potential of *fd*-Y21M rods dissolved in the background membrane is given by $\mu = k_B T \ln(C_{\text{BG}}A_{\text{rod}})$. The net change in free energy associated with taking a single *fd*-Y21M virus from the background and inserting it into a raft containing N rods is given by $\Delta G - \mu$ (Fig. 2i, inset). The absolute raft free-energy landscape up to a constant offset was obtained by numerically integrating $\Delta G - \mu$ with the use of a cumulative trapezoidal rule (Fig. 2i).

Single-molecule analysis. We prepared raft-bearing membranes with non-fluorescent long and short rods. These membranes were doped with a very low volume fraction of fluorescent short rods ($\sim 0.003\%$) enabling us to visualize the motion of single rods within the membrane background and rafts. The membranes were simultaneously visualized in phase contrast and fluorescence microscopy to determine the number of raft-bound labelled rods, N_f . Standard analysis shows that the decay of the autocorrelation function of N_f is given by $\langle N_f(t)N_f(t+\Delta t) \rangle - \langle N_f(t) \rangle^2 = (k_{\text{on}}C_{\text{BG}}/k_{\text{off}})e^{-k_{\text{off}}\Delta t}$ (ref. 37). Therefore k_{off} and k_{on} can be determined uniquely with this expression. These measurements were performed for equilibrium-sized rafts only.

Measurement of membrane-mediated repulsive raft interactions. To measure effective raft–raft interactions, we used the blinking optical trap technique originally developed for measuring effective potentials of conventional spherical colloids²⁴. Being enriched in shorter *fd*-Y21M rods, rafts are repelled from an optical trap. To manipulate these rafts, we created an optical plough-like configuration consisting of multiple time-shared optical traps generated by an acousto-optic deflector. Using a pair of optical ploughs we brought two rafts into close proximity (Supplementary Video 5). Once the traps were switched off, the repulsive membrane-mediated interactions drove the rafts apart. It was possible to measure these interactions by repeating the blinking experiment dozens of times and analysing the time evolution of subsequent raft trajectories. Focusing an optical trap onto a membrane induces local distortions that could potentially affect the measurement of raft interactions. This made it advantageous to measure interactions using blinking optical traps. In this technique, the traps are only used to bring two rafts into an initial low-probability configuration, but the actual measurement is performed when the optical traps are switched off. On switching the laser off, the trap-induced membrane distortions

relaxed on a millisecond timescale, which is essentially instantaneous when compared with the timescale of minutes over which two repulsive rafts drift apart.

Optical plough trap configurations were generated by time sharing a laser beam (4W, 1064 nm, Compass 1064, Coherent) using a pair of orthogonally oriented paratellurite acousto-optic deflectors (Intra-Action). The laser beam was projected onto the back focal plane of an oil-immersion objective (1.3 numerical aperture, $\times 100$ PlanFluor) and focused onto the imaging plane. The multiple trap locations were specified with a custom LABVIEW software. Raft separations were measured as a function of time, once the traps had been switched off, using standard video tracking methods³⁸. The time lapse between successive frames was 100 ms and the exposure time was 50 ms.

The discretized probability of a raft pair's being separated by distance r_j at time $t + \Delta t$ is given by

$$\rho(r_j) = \sum_i P(r_i \rightarrow r_j) \rho(r_i)$$

where P is the transition probability for a raft pair initially separated by distance r_i at time t to be separated by distance r_j at time $t + \Delta t$ later. $\rho(r_i)$ is the probability of a raft pair's being separated by a distance r_i at time t (ref. 22). Experimentally, P is

determined by binning the trajectories according to the initial and final separations in each time step. The steady-state solution to equation (1) is equal to the equilibrium probability, $\rho(r)$, computed by calculating the eigenvector of the transition probability matrix. The inter-raft potential is given by $U(r) = -k_B T \ln(\rho(r))$.

31. Dogic, Z. & Fraden, S. Development of model colloidal liquid crystals and the kinetics of the isotropic–smectic transition. *Phil. Trans. R. Soc. A* **359**, 997–1014 (2001).
32. Maniatis, T., Sambrook, J. & Fritsch, E. *Molecular Cloning* (Cold Spring Harbor Laboratory, 1989).
33. Lettinga, M. P., Barry, E. & Dogic, Z. Self-diffusion of rod-like viruses in the nematic phase. *Europhys. Lett.* **71**, 692–698 (2005).
34. Lau, A. W. C., Prasad, A. & Dogic, Z. Condensation of isolated semi-flexible filaments driven by depletion interactions. *Europhys. Lett.* **87**, 48006 (2009).
35. Oldenbourg, R. & Mei, G. New polarized-light microscope with precision universal compensator. *J. Microsc.* **180**, 140–147 (1995).
36. Zakhary, M. J. *et al.* Imprintable membranes from incomplete chiral coalescence. *Nature Commun.* **5**, <http://dx.doi.org/10.1038/ncomms4063> (2014).
37. Gardiner, C. W. *Handbook of Stochastic Methods* (Springer, 1985).
38. Crocker, J. C. & Grier, D. G. Methods of digital video microscopy for colloidal studies. *J. Colloid Interf. Sci.* **179**, 298–310 (1996).

Temperature sensitivity of soil respiration rates enhanced by microbial community response

Kristiina Karhu^{1†}, Marc D. Auffret², Jennifer A. J. Dungait³, David W. Hopkins⁴, James I. Prosser², Brajesh K. Singh⁵, Jens-Arne Subke⁶, Philip A. Wookey⁷, Göran I. Ågren⁸, Maria-Teresa Sebastià^{9,10}, Fabrice Gouriveau⁹, Göran Bergkvist¹¹, Patrick Meir^{12,13}, Andrew T. Nottingham¹², Norma Salinas¹⁴ & Iain P. Hartley¹

Soils store about four times as much carbon as plant biomass¹, and soil microbial respiration releases about 60 petagrams of carbon per year to the atmosphere as carbon dioxide². Short-term experiments have shown that soil microbial respiration increases exponentially with temperature³. This information has been incorporated into soil carbon and Earth-system models, which suggest that warming-induced increases in carbon dioxide release from soils represent an important positive feedback loop that could influence twenty-first-century climate change⁴. The magnitude of this feedback remains uncertain, however, not least because the response of soil microbial communities to changing temperatures has the potential to either decrease^{5–7} or increase^{8,9} warming-induced carbon losses substantially. Here we collect soils from different ecosystems along a climate gradient from the Arctic to the Amazon and investigate how microbial community-level responses control the temperature sensitivity of soil respiration. We find that the microbial community-level response more often enhances than reduces the mid- to long-term (90 days) temperature sensitivity of respiration. Furthermore, the strongest enhancing responses were observed in soils with high carbon-to-nitrogen ratios and in soils from cold climatic regions. After 90 days, microbial community responses increased the temperature sensitivity of respiration in high-latitude soils by a factor of 1.4 compared to the instantaneous temperature response. This suggests that the substantial carbon stores in Arctic and boreal soils could be more vulnerable to climate warming than currently predicted.

Short-term experiments have demonstrated that the rate of soil microbial respiration increases exponentially with temperature, and this general relationship has been used to parameterize soil carbon (C) and Earth-system models^{4,10}. However, plant physiologists have demonstrated that short-term measurements are inadequate for representing the dynamic response of plant respiration to changes in temperature. In plants, thermal acclimation, defined as the “subsequent adjustment in the rate of respiration to compensate for an initial change in temperature”¹¹, greatly reduces the impact of temperature changes on respiration in the medium- to long-term, and incorporating this acclimation into models alters predicted rates of terrestrial C uptake¹². In soil, there may be a response in microbial communities that is analogous to thermal acclimation in plants, given that these communities adapt to changes in temperature¹³. However, it is unclear whether microbial community responses always reduce the effect of a temperature change on respiration rates. Responses that enhance the instantaneous effect of temperature changes on soil respiration have also been observed^{8,9,14}. Until now there has been no large-scale evaluation of the role of microbial community responses in controlling the temperature sensitivity of soil

respiration, adding considerable uncertainty to predictions of the magnitude and direction of C-cycle feedbacks to climate change¹⁵.

Despite several attempts at clarification, the use of terminology remains quite confused in this research field¹³. Because measurements of soil microbial respiration are made at the level of the whole community, they encompass acclimation (physiological responses of individuals), adaptation (genetic changes within species) and ecological responses (for example, competition altering species composition), all of which can lead to adjustments in respiration rates following a sustained change in temperature¹³. For this reason, rather than acclimation or adaptation, which have strict definitions, we have chosen to use the term ‘community-level response’. These community-level responses can be either compensatory or enhancing (that is, reducing or increasing the effect of a temperature change on respiration rates in the longer term). We investigated how microbial community-level responses affect the temperature sensitivity of soil respiration.

When soil is warmed for an extended period, the initial increase in biological activity leads to a loss of readily decomposable C (ref. 5). Microbial activity then tends to decline in the longer term, but it is often impossible to determine whether this is caused by the loss of the readily decomposable C or by a compensatory response of the microbial community, given that both would reduce activity^{16,17}. To differentiate between these two mechanisms, we established an approach⁸ that involves cooling soil in the laboratory. Compensatory community responses and substrate loss should have opposite effects on microbial activity under cooling conditions. In the absence of C inputs, soil C losses still occur in cooled soils, thus reducing activity, albeit at a slower rate than in the controls. However, a compensatory response of the microbial community should result in a gradual increase in respiration rate as the community compensates for the effects of the cooling; this is analogous to what is observed for thermal acclimation of plant respiration¹¹. Furthermore, because we can quantify rates of soil C loss, we can also identify enhancing responses if respiration rates decline more rapidly in the cooled soil than in the control.

Using our cooling approach, we carried out a global investigation of how microbial community responses to temperature changes affect soil respiration rates, collecting soil from sites representing a range of ecosystem types (arable, grassland, deciduous and evergreen broadleaf forest, coniferous forest and heath) across a gradient of mean annual temperature (MAT) from -6°C to 24°C (Fig. 1 and Extended Data Table 1). Twenty samples of each soil were pre-incubated at 3°C above the MAT of their collection site (see Fig. 2a) for 84 days to allow respiration rates to stabilize. On day 84, five samples were destructively sampled for microbial biomass determination, ten samples were cooled

¹Geography, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4RJ, UK. ²Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen AB24 3UU, UK.

³Rothamsted Research—North Wyke, Okehampton, Devon EX20 2SB, UK. ⁴School of Agriculture, Food & Environment, The Royal Agricultural University, Cirencester, Gloucestershire GL7 6JS, UK.

⁵Hawkesbury Institute for the Environment, University of Western Sydney, Penrith 2751, New South Wales, Australia. ⁶School of Natural Sciences, Biological and Environmental Sciences, University of

Stirling, Stirling FK9 4LA, UK. ⁷School of Life Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK. ⁸Department of Ecology, Swedish University of Agricultural Sciences (SLU), 750 07 Uppsala,

Sweden. ⁹Laboratory of Functional Ecology and Global Change, Forest Sciences Centre of Catalonia (CTFC), 25280 Solsona, Spain. ¹⁰Department of Horticulture, Botany and Landscaping, School of

Agrifood and Forestry Science and Engineering, University of Lleida, 25198 Lleida, Spain. ¹¹Department of Crop Production Ecology, Swedish University of Agricultural Sciences (SLU), 750 07 Uppsala,

Sweden. ¹²School of Geosciences, University of Edinburgh, Edinburgh EH8 9XP, UK. ¹³Research School of Biology, The Australian National University, Canberra, Australian Capital Territory 0200, Australia.

¹⁴Sección Química, Pontificia Universidad Católica del Perú, Lima 32, Perú. [†]Present address: Department of Forest Sciences, University of Helsinki, 00014 Helsinki, Finland.

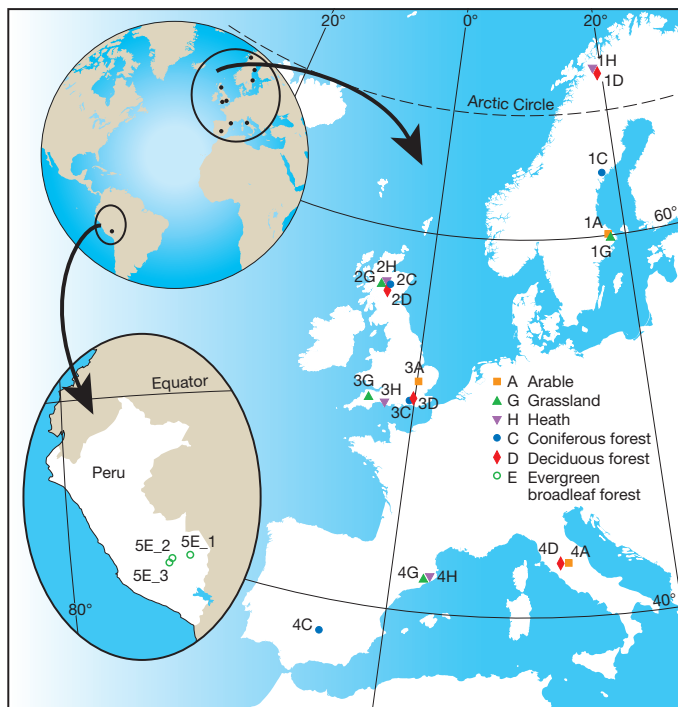


Figure 1 | Soil was sampled from boreal and Arctic, temperate, Mediterranean and tropical climates. Arable, grassland, heath, coniferous forest and deciduous forest (A, G, H, C and D) sites were sampled in each climatic region (except the tropics, where evergreen (E) broadleaf forest sites were sampled along an altitudinal gradient in the Peruvian Andes), and within each ecosystem type, sites are numbered from 1 to 5 in order of increasing MAT. Details of sampling sites (vegetation and soil characteristics) are presented in Extended Data Table 1.

by 6 °C (MAT minus 3 °C), and five controls were maintained at MAT plus 3 °C for the remaining 90 days of the experiment. Five of the cooled samples were incubated at MAT minus 3 °C for 90 days, a time period relevant to seasonal changes in temperature, which have been hypothesized to cause thermal adaptation¹⁸. The other five cooled samples were rewarmed to MAT plus 3 °C after 60 days at MAT minus 3 °C, and incubated at MAT plus 3 °C for the remaining 30 days of the experiment, allowing the reversibility of any response to be determined.

Our approach establishes two clear criteria for quantifying either compensatory or enhancing community-level responses (Fig. 2a and Extended Data Fig. 1). First, the carbon dioxide (CO₂) flux, normalized to the flux at the time of cooling (control samples) or immediately after cooling (cooled samples), was plotted against cumulative C loss (see Methods and Extended Data Figs 1–7). The impact of community

responses on respiration rates at the measurement temperatures (RR_{MT} , where RR represents the response ratio) was calculated as the normalized control respiration rate, at the percentage C loss corresponding to the total percentage C loss in the cooled soils (see Supplementary Information), divided by the normalized cooled respiration rate at the end of the incubation. Ratios <1 indicate a compensatory response (that is, normalized respiration rates were greater at a given level of soil C loss in the cooled treatment), and ratios >1 indicate an enhancing response (that is, normalized respiration rates were lower at a given level of soil C loss in the cooled treatment). A second quantitative measure was obtained by comparing the respiration rates of samples rewarmed after 60 days of cooling with control sample respiration rates at the same C loss (see Supplementary Information). This ratio at a common temperature¹⁹ (RR_{CT}) was calculated as the control respiration rate divided by rewarmed respiration rate and, again, ratios <1 and >1 indicate compensatory and enhancing responses, respectively. Given that changes in biomass have been considered to be important in previous studies⁸, we also calculated RR_{MT} on a microbial-biomass-specific basis (see Methods).

All three possible community-level responses were observed: compensatory responses (Fig. 2c), enhancing responses (Fig. 2d) and no response (Fig. 2b). However, for the 22 soils analysed, many more statistically significant cases of enhancing responses were observed (see Supplementary Information). Overall average response ratios ($n = 22$ soils) were significantly above 1 ($P < 0.01$ for RR_{MT} , Fig. 3a; $P < 0.05$ for RR_{CT} , Fig. 3b). In all cases of clear enhancing or compensatory responses, respiration rates after rewarming subsequently approached control rates (see Fig. 2c, d). This reversibility of the response indicates that the patterns were not caused by cooling altering the decomposability of the remaining C, and emphasizes the comparability, in terms of effects on rates of respiration, of microbial community responses to cooling and warming.

The average RR_{MT} values were greatest for boreal and Arctic soils (MAT < 7 °C group; Fig. 3a), but also significantly above 1 for the MAT > 14 °C group. For the MAT < 7 °C group, the microbial community response increased the temperature sensitivity of soil respiration by a factor of 1.4 during the 90 days of cooling; the temperature sensitivity, expressed as a Q_{10} value (proportional change in respiration for a 10 °C change in temperature), increased from 4.6 at the time of cooling to 6.3 at the end of the incubation (see Methods).

Arable or ‘managed’, low-C-content, and low carbon-to-nitrogen ratio (C:N) soils were the only soils to show average RR_{MT} values close to or below 1 (Fig. 3a). Enhancing responses were generally more common in soils with high C content, high C:N ratios and low pH values (Fig. 3a); forest and ‘natural ecosystem’ groups also showed enhancing responses. C:N ratio was the only soil or site variable (Fig. 3a) that was significantly correlated with the RR_{MT} responses across all data ($\ln[RR_{MT}] = 0.188 \times \ln[C:N] - 0.406$, $R^2 = 0.335$, $P = 0.005$), and the low RR_{MT} for the MAT 7 °C–14 °C group may have been related to the greater

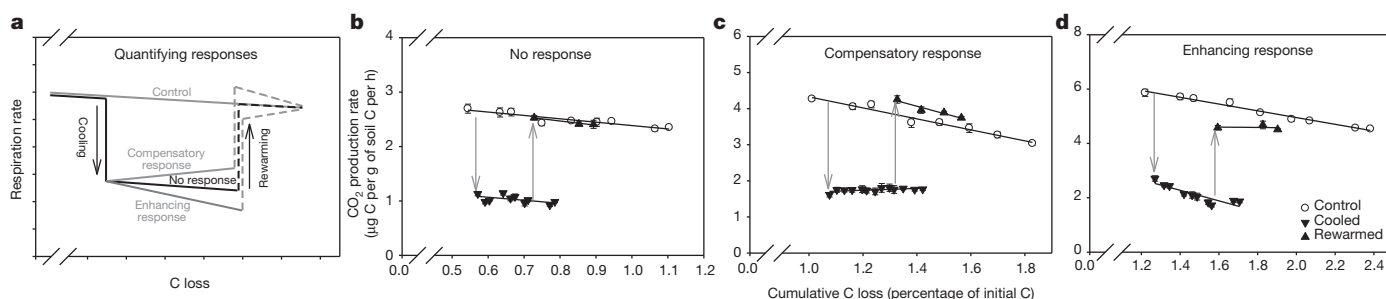


Figure 2 | The patterns of CO₂ flux that would be observed in the case of no response, compensatory and enhancing community-level responses. The schematic diagram (a), indicates how a gradual increase in soil respiration rate after cooling provides support for a compensatory response, while a more rapid decline in cooled soils indicates an enhancing response, as well as how differences in rates of respiration in rewarmed versus control samples can be

used to quantify the magnitude and direction of the community-level response (see also Extended Data Fig. 1). **b–d**, Examples of measured CO₂ fluxes illustrating no response (b, soil 2C), compensatory response (c, soil 3A) and enhancing response (d, soil 1C) given as mean respiration rates ± 1 standard error ($n = 5$). A break in the x-axis scale denotes that pre-incubation data are not shown.

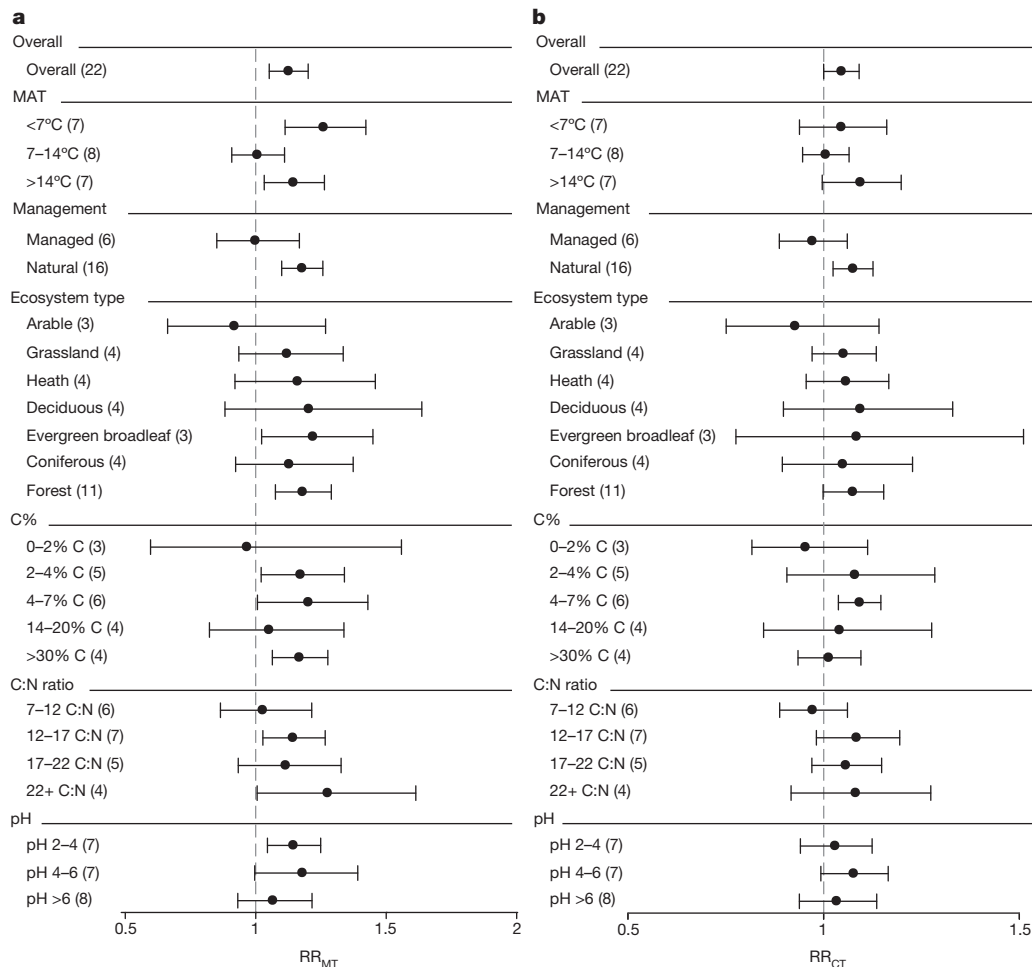


Figure 3 | The impact of the microbial community responses on the response of soil respiration to changes in temperature. The mean \pm 95% confidence intervals of RR_{MT} (a) and RR_{CT} values (b) are presented overall (that is, including all data), and for different soil groups, based on ecosystem

type, management, climate and various soil properties (n is given in parentheses). Values >1 indicate an enhancing response, and values <1 indicate a compensatory response.

number of managed, low C:N ratio soils in this group rather than the temperature range itself.

Overall, our results demonstrate that microbial-community-level responses enhance the impacts of temperature changes on soil respiration rates. To improve mechanistic understanding, and for modelling dynamics^{15,20}, it has been argued that changes in biomass must be accounted for explicitly in quantifying microbial community responses¹³. In our study, for a given C loss, biomass did not differ much between cooled and control soils (whether measured by chloroform-fumigation extraction or quantitative polymerase chain reaction, qPCR), and thus mass-specific patterns did not differ substantially from the raw responses; mass-specific responses (RR_{MT_MS}) had slightly greater average values ($RR_{MT_MS} > RR_{MT}$) but also showed greater variability (Extended Data Fig. 8 and Supplementary Information). Overall, changes in microbial biomass could not explain the observed microbial community responses.

The greater enhancing responses in cold soils and in soils with high C:N ratios require further consideration. The requirements for surviving at low temperatures are known to present strong selection pressures that induce fundamental changes at the cellular level^{21,22}. In plants, cold acclimation results in an upregulation of respiration rates at lower temperatures¹¹, but our data demonstrate that adaptation to cooling by microbial communities in high-latitude soils reduces respiration rates, which may be consistent with strategies that promote survival but reduce metabolic activity^{21,22}. However, strong enhancing responses were also observed in some tropical and Mediterranean soils, so the development of cold tolerance cannot be the full explanation for the observed

responses. C:N was the only variable that was positively correlated with RR_{MT} across all data. If the temperature sensitivities of key N-cycle processes are greater than some C-cycle processes^{23,24}, then it is possible that N availability may limit microbial activity following cooling, especially in soils with high C:N values. This could potentially induce adaptive changes in allocation to N versus C acquisition to meet stoichiometric requirements²³, which could in turn be reversed on rewarming. Links between C and N cycling may also help to explain why our results differ from some previous studies. Compensatory thermal adaptation has previously been observed in ectomycorrhizal fungi grown on agar¹⁸, and also in monocultures of heterotrophic fungi⁷. The dominance of enhancing adaptation responses identified in our study could be related to the fact that community-level competition for C and N sources is important for determining the overall response to warming.

In conclusion, enhancing community-level responses were much more common than compensatory responses, with the latter mainly limited to arable soils and soils with low C content (Fig. 3), thus limiting the potential importance of compensatory responses for rates of climate-change-induced C losses. The predominance of enhancing responses implies that decreased soil respiration rates in response to long-term ecosystem warming in the field²⁵ are probably related to the loss of readily decomposable C, rather than to any community-level response downregulating microbial respiration rates. Finally, given that boreal and arctic regions contain more than half of the global soil C stock²⁶, the strong enhancing responses observed in these soils could have important consequences for the global C budget.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 February; accepted 17 June 2014.

1. Stocker, T. F. *et al.* (eds) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2013).
2. Shao, P., Zeng, Z., Moore, D. J. P. & Zeng, X. Soil microbial respiration from observations and earth system models. *Environ. Res. Lett.* **8**, 034034 (2013).
3. Davidson, E. A. & Janssens, I. A. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* **440**, 165–173 (2006).
4. Friedlingstein, P. *et al.* Climate-carbon cycle feedback analysis: results from the C4MIP model intercomparison. *J. Clim.* **19**, 3337–3353 (2006).
5. Bradford, M. A. *et al.* Thermal adaptation of soil microbial respiration to elevated temperature. *Ecol. Lett.* **11**, 1316–1327 (2008).
6. Bradford, M. A., Watts, B. W. & Davies, C. A. Thermal adaptation of heterotrophic soil respiration in laboratory microcosms. *Glob. Change Biol.* **16**, 1576–1588 (2010).
7. Crowther, T. W. & Bradford, M. A. Thermal acclimation in widespread heterotrophic soil microbes. *Ecol. Lett.* **16**, 469–477 (2013).
8. Hartley, I. P. *et al.* Soil microbial respiration in arctic soil does not acclimate to temperature. *Ecol. Lett.* **11**, 1092–1100 (2008).
9. Nie, M. *et al.* Positive climate feedbacks of soil microbial communities in a semi-arid grassland. *Ecol. Lett.* **16**, 234–241 (2013).
10. Kirschbaum, M. U. F. The temperature dependence of organic-matter decomposition—still a topic of debate. *Soil Biol. Biochem.* **38**, 2510–2518 (2006).
11. Atkin, O. K. & Tjoelker, M. G. Thermal acclimation and the dynamic response of plant respiration to temperature. *Trends Plant Sci.* **8**, 343–351 (2003).
12. Galbraith, D. *et al.* Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change. *New Phytol.* **187**, 647–665 (2010).
13. Bradford, M. A. Thermal adaptation of decomposer communities in warming soils. *Front. Microbiol.* **4**, 00333 (2013).
14. Hartley, I. P., Heinemayer, A., Evans, S. P. & Ineson, P. The effect of soil warming on bulk soil vs. rhizosphere respiration. *Glob. Change Biol.* **13**, 2654–2667 (2007).
15. Wieder, W. R., Bonan, G. B. & Allison, S. D. Global soil carbon projections are improved by modelling microbial processes. *Nature Clim. Change* **3**, 909–912 (2013).
16. Ågren, G. I. & Bosatta, E. Reconciling differences in predictions of temperature response of soil organic matter. *Soil Biol. Biochem.* **34**, 129–132 (2002).
17. Kirschbaum, M. U. F. Soil respiration under prolonged soil warming: are rate reductions caused by acclimation or substrate loss? *Glob. Change Biol.* **10**, 1870–1877 (2004).
18. Malcolm, G. M., López-Gutiérrez, J. C., Koide, R. T. & Eissenstat, D. M. Acclimation to temperature and temperature sensitivity of metabolism by ectomycorrhizal fungi. *Glob. Change Biol.* **14**, 1169–1180 (2008).
19. Loveys, B. R. *et al.* Thermal acclimation of leaf and root respiration: an investigation comparing inherently fast- and slow-growing plant species. *Glob. Change Biol.* **9**, 895–910 (2003).
20. Allison, S. D., Wallenstein, M. D. & Bradford, M. A. Soil-carbon response to warming dependent on microbial physiology. *Nature Geosci.* **3**, 336–340 (2010).
21. Hochachka, P. W. & Somero, G. N. *Biochemical Adaptation: Mechanism and Process in Physiological Evolution* 466 (Oxford Univ. Press, 2002).
22. Schimel, J., Balser, T. C. & Wallenstein, M. Microbial stress-response physiology and its implications for ecosystem function. *Ecology* **88**, 1386–1394 (2007).
23. Billings, S. A. & Ballantyne, F. IV. How interactions between microbial resource demands, soil organic matter stoichiometry, and substrate reactivity determine the direction and magnitude of soil respiratory responses to warming. *Glob. Change Biol.* **19**, 90–102 (2013).
24. Lehmeier, C. A., Min, K., Niehues, N. D., Ballantyne, F., IV & Billings, S. A. Temperature-mediated changes of exoenzyme-substrate reaction rates and their consequences for carbon to nitrogen flow ratio of liberated resources. *Soil Biol. Biochem.* **57**, 374–382 (2013).
25. Luo, Y., Wan, S., Hui, D. & Wallace, L. L. Acclimatization of soil respiration to warming in a tall grass prairie. *Nature* **413**, 622–625 (2001).
26. FAO, IIASA, ISRIC, ISSCAS & JRC. *Harmonized World Soil Database Version 1.2*, <http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/> (FAO and IIASA, 2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the staff of the Forestry Commission at Alice Holt Forest, T. Taylor from RSPB Aylesbeare Common Reserve, J. Harris from Cranfield University, C. Moscatelli and S. Marinari from Tuscia University, J. A. Carreira de la Fuente from the University of Jaén, R. Giesler from Umeå University and E. Cosio from The Pontifical Catholic University of Peru for help with site selection and soil sampling. We thank N. England for technical assistance with constructing the incubation system, J. Zaragoza Castells for help with soil sampling, A. Elliot for conducting the particle size analyses, J. Grapes for help with carbon and nitrogen analysis and S. Rouillard, H. Jones and T. Kurtén for assistance with graphics. This work was carried out with Natural Environment Research Council (NERC) funding (grant number NE/H022333/1). K.K. was supported by an Academy of Finland post-doctoral research grant while finalizing this manuscript. P.M. was supported by ARC FT110100457 and NERC NE/G018278/1, and B.K.S. by the Grain Research and Development Corporation and ARC DP130104841.

Author Contributions K.K. conducted the CO₂ measurements and statistical analyses. K.K. and M.D.A. conducted the chloroform-fumigation extraction and qPCR analyses, respectively, and led the data analysis and interpretation. I.P.H. (lead investigator), P.A.W., D.W.H., B.K.S. and J.I.P. designed the study. G.I.Å. and K.K. were responsible for the modelling presented in the methods. K.K., I.P.H., J.A.J.D., D.W.H., J.-A.S., P.A.W., M.-T.S., F.G., G.B., P.M., A.T.N. and N.S. were involved in planning site selection and soil sampling. All authors were involved in interpreting the results and contributed to writing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.K. (kristiina.karhu@helsinki.fi).

Asymmetric three-dimensional topography over mantle plumes

Evgueni Burov^{1,2} & Taras Gerya³

The role of mantle–lithosphere interactions in shaping surface topography has long been debated^{1–3}. In general^{3,4}, it is supposed that mantle plumes and vertical mantle flows result in axisymmetric, long-wavelength topography, which strongly differs from the generally asymmetric short-wavelength topography created by intraplate tectonic forces. However, identification of mantle-induced topography is difficult³, especially in the continents⁵. It can be argued therefore that complex brittle–ductile rheology and stratification of the continental lithosphere result in short-wavelength modulation and localization of deformation induced by mantle flow⁶. This deformation should also be affected by far-field stresses and, hence, interplay with the ‘tectonic’ topography (for example, in the ‘active/passive’ rifting scenario^{7,8}). Testing these ideas requires fully coupled three-dimensional numerical modelling of mantle–lithosphere interactions, which so far has not been possible owing to the conceptual and technical limitations of earlier approaches. Here we present new, ultra-high-resolution, three-dimensional numerical experiments on topography over mantle plumes, incorporating a weakly pre-stressed (ultra-slow spreading), rheologically realistic lithosphere. The results show complex surface evolution, which is very different from the smooth, radially symmetric patterns usually assumed as the canonical surface signature of mantle upwellings⁹. In particular, the topography exhibits strongly asymmetric, small-scale, three-dimensional features, which include narrow and wide rifts, flexural flank uplifts and fault structures. This suggests a dominant role for continental rheological structure and intra-plate stresses in controlling dynamic topography, mantle–lithosphere interactions, and continental break-up processes above mantle plumes.

The question of the spatial scales at which surface topography is affected by mantle–lithosphere interactions rather than by plate-scale processes has attracted attention^{1–3,5–10} owing to its key importance for understanding the Earth’s tectonic, magmatic and seismic activity^{4,8}. This problem becomes crucial specifically in the light of current debates concerning the existence and impact of mantle plumes¹¹, the significance of the LAB (lithosphere–asthenosphere boundary), large-scale mantle–lithosphere instabilities¹² and lithosphere rheology¹³, and the mechanisms of continental rifting and breakup¹⁴. Of particular interest is the asymmetric or complex surface topography in regions where plate mechanics might influence mantle dynamics, such as the East African^{2,8,14,15}, Tanzanian^{8,14,15}, Red Sea^{8,14} and Dead Sea rift systems^{8,14}, the Pannonian–Carpathian system¹⁵, the Basin and Range province¹⁴, the Rio Grande rift and the Colorado plateau¹⁴, the Iranian and Tibetan plateaus¹⁴, the Yellowstone hotspot¹⁵, and some cratons^{15,16}. Mantle–lithosphere interactions have important consequences for tectonic evolution and for long-term climate change. For example, Archaean metallogenic ‘crises’ (sharp accelerations of metallogenic production) at the boundaries of the West African and Australian cratons and formation of volcanic traps coincide with plume events^{15–17} that have also played an important role in continental crustal accretion¹⁸ and in key geodynamic and environmental processes; and Siberian Traps magmatism is behind the greenhouse-gas emissions that led to global climate catastrophe and mass extinction of 80% of living species at the end of the Permian¹⁷.

The observation-based contribution of mantle flow to surface topography, dubbed ‘dynamic topography’³, is commonly evaluated by filtering out short-wavelength components ($\lambda < 1,000$ km) and by stripping local isostatic components of the observed topography and gravity and geoid fields². The remaining long-wavelength ($\lambda > 1,000$ km) undulations with amplitudes of the order of several hundred metres are regarded as an imprint of mantle flow. Such undulations are however difficult to separate from contributions due to heterogeneities in lithospheric structure, surface erosion and, in some cases, tectonic processes. Mantle flow is only hinted at by other observations such as gravity, seismic tomography, seismic anisotropy patterns (SKS orientations) or tectonic strain orientations^{2,3,10,19–23}. For instance, seismic tomography, which has been used to characterize mantle flow, does not always reveal features such as rising mantle plumes and sinking lithospheric slabs¹⁹. As a consequence, attempts to link mantle flow with surface dynamics often yield ambiguous results^{3,9}. These complications have triggered debates on the existence of mantle plumes^{11,19,21}, partitioning between mantle and tectonically driven processes^{6,7,9,13–15} and other questions such as the significance of dynamic topography^{9,20}. These disputes are fuelled, in particular, by the fact that axisymmetric dome uplifts expected above mantle upwellings^{1,6} are not unambiguously distinguished in the areas showing other (for example, tomographic, thermal, geochemical) evidence for mantle upwellings⁵: such areas include major recognized hotspots such as the African Superplume^{1,2}, Hawaii^{24,25}, and the ‘Afar Triangle’–Red Sea zone^{26,27}.

Surface impact of mantle–lithosphere interactions is usually predicted from three-dimensional (3D) numerical models of mantle dynamics^{2,9,10}. However, the existing models^{2,9,10} are designed to reproduce mantle flow and not surface deformation, as they: (1) do not compute surface topography (estimated *post hoc* from isostatic assumptions, Fig. 1a); (2) imply single-layer viscous stagnant lithosphere; and (3) have low spatial resolution, insufficient to resolve small-scale deformation and rheological stratification¹³.

High-resolution two-dimensional (2D) models^{6,15,28,29} of rheologically stratified lithosphere that are free of these limitations predict short-wavelength topography marked by closely spaced fault patterns and multi-harmonic tectonic-scale features with $\lambda \approx 30$ –100, 150–350 and 400–500 km. The reason for this is that thick rheologically stratified lithosphere acts to dampen and modulate the mantle flow ‘signal’ (Fig. 1). This is because the lower continental crust typically represents a low-viscosity 10–20-km-thick ductile channel^{6,13} that mechanically separates crust from mantle and absorbs undulations of crust–mantle interface (Moho) induced by mantle flow (Fig. 1). As a consequence, mantle flow stresses acting at the LAB are attenuated or re-focused, resulting in short-wavelength tensional/compressional instabilities in competent rheological layers constituting lithosphere. The actual wavelengths are controlled by the thicknesses of these layers and do not depend on the wavelength of mantle upwellings²⁹.

The 2D models provide little information, however, on 3D features of Earth’s surface topography. This is despite the fact that axisymmetric mantle upwellings or downwellings would be expected to produce axisymmetric radial topography and gravity, which are therefore searched-for in

¹Sorbonne Universités, UPMC Université Paris 06, UMR 7193, Institut des Sciences de la Terre Paris (ISTeP), F-75005 Paris, France. ²CNRS, UMR 7193, Institut des Sciences de la Terre Paris (ISTeP), F-75005 Paris, France. ³ETH, 8092 Zürich, Switzerland.

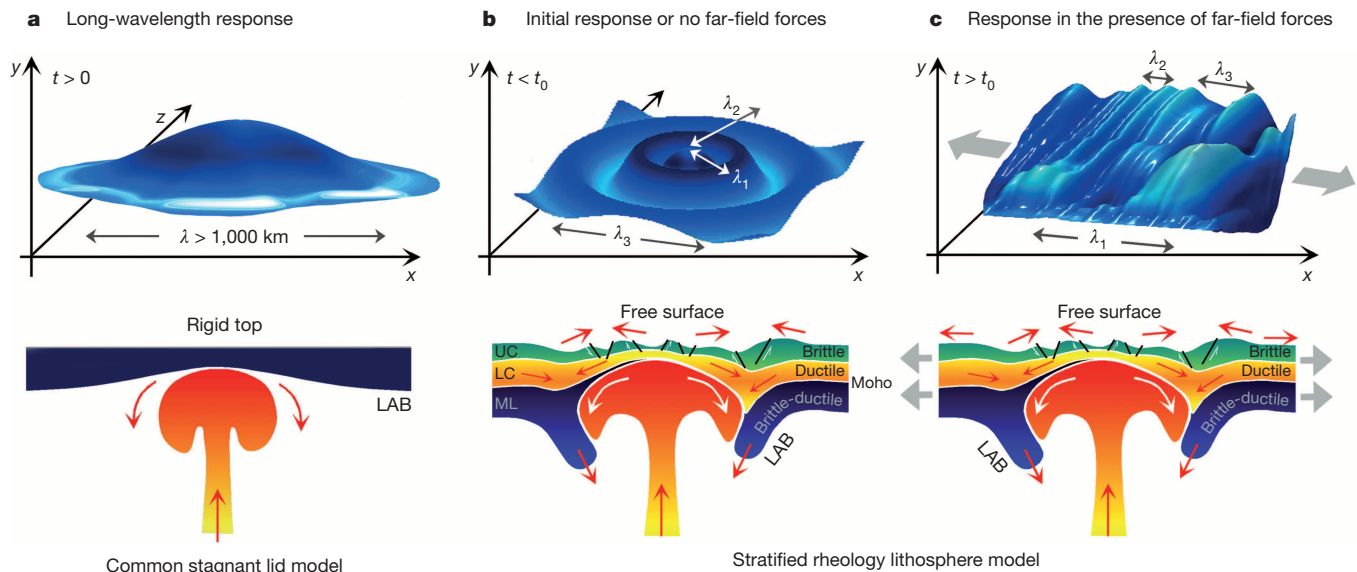


Figure 1 | Mantle–lithosphere interactions and topography. Evaluation of deformation of the Earth's surface caused by mantle flow (dynamic topography³) is commonly not straightforward and is based on additional assumptions; here we show surface topography (top row) and corresponding conceptual models (bottom row). In the conventional^{11,3,9} view (a), mantle upwellings can produce only long-wavelength ($\lambda > 1,000$ km) axisymmetric domes at the surface; in the corresponding models, the lithosphere commonly represents a passive flat-top viscous stagnant lid^{2,9,10} and the topography is computed from mantle flow patterns assuming local isostasy¹. Our approach (b, c) considers instead free-surface response of rheologically stratified

brittle–ductile lithosphere that may exhibit several short-wavelength harmonics^{6,29}, whereas long-wavelength deformation is dampened in the ductile lower crust (LC). The actual surface wavelengths λ_i (for example, $\lambda_1 \approx 30$ –50, $\lambda_2 \approx 100$ –150, $\lambda_3 \approx 250$ –350 km) are controlled by thicknesses of strong upper crustal (UC), lower crustal (LC) and mantle–lithosphere (ML) layers⁶. The initial topography (b) may become strongly asymmetric after some short time t_0 in the case of even very weak far-field tectonic stresses that trigger anisotropic strain localizations (c). In the bottom row, red arrows and white arrows indicate direction of surface and subsurface movements. Black lines indicate faults, and white half-arrows indicate direction of movements on the fault interfaces.

spectral studies attempting to identify dynamic topography^{5,15} (Fig. 1). Yet, in the presence of a directional far stress/strain field, mantle flow may trigger development of non-axisymmetric features that are controlled by far-field stress/strain orientation and strain-localizing properties of the pre-existing lithosphere⁸ (Methods, Fig. 1).

These considerations suggest a need for a novel, tectonically realistic 3D modelling approach that encompasses large horizontal scales ($>1,000$ km \times 1,000 km), the entire 650-km-deep upper mantle, and offers 'lithospheric-grade' numerical resolution ($\sim 2 \times 2 \times 2$ km) in order to handle rheological stratification and localized deformation²⁹. This implies unprecedented numerical efforts, huge mesh dimensions comprising several hundred million elements ($\sim 500^3$), and billions of moving rock markers, compared to the previous models that could handle only 10 times smaller arrays. To meet this challenge, we have designed an ultra-high-resolution model based on an optimized staggered grid/particle-in-cell method³⁰ (Methods). Several sets of 3D experiments exploring different geodynamic settings have been performed, totalling 10^6 hours (~ 114 years) of nodal processing time on the supercomputers of ETH and UPMC.

To reduce the tested parameter range, we conducted a series of fast 2D high-resolution parametric experiments^{6,15,29} (Methods, Extended Data Figs 1–4). The retained reference model considers a 200-km-wide plume impinging on the base of a 150-km-thick, 250-Myr-old lithosphere²⁹. We then conducted a set of 3D experiments for different far-field tectonic boundary conditions: (1) no far-field forces; (2) unidirectional pure shear extensional forcing; (3) bi-directional pure shear forcing; and (4) combined pure and simple shear forcing. The lateral boundary conditions correspond to ultra-slow velocities applied at model sides (3 mm yr^{-1} , typical for pre-breakup continental rifts affected by mantle upwellings, such as the Rio Grande and East African rifts¹⁴), so that the associated intra-plate stresses are very low—insufficient on their own to produce significant large-scale deformation (for example, passive rifting) during the time of plume impingement.

Figure 2 illustrates two end-member scenarios: ultra-slow stretching of a continental lithosphere in the absence of active mantle upwelling;

and active mantle plume upwelling impinging on a stress-free lithosphere. In the first case (Fig. 2a), ultra-slow tectonic stretching produces patterns of distributed small-offset parallel faults, in which the fault spacing (30–50 km) is controlled by the thickness of the upper-crustal brittle layer. These faults are distributed over 1,000-km scales, without focusing in any particular zone. In the second scenario (Fig. 2b), surface deformation is radially axisymmetric, with periodic multi-harmonic short-wavelength radial undulations ($\lambda \approx 30$ –250 km, Fig. 2b, Extended Data Fig. 5). This result is consistent with rheologically realistic 2D models^{6,15,29} (Extended Data Figs 1–4), and (expectedly) disagrees with conventional 3D models^{9,10} ($\lambda > 1,000$ km). We note radial undulations at the bottom (LAB) of the lithosphere, with $\lambda \approx 200$ –250 km, and shorter-wavelength ($\lambda \approx 30$ –100 km) surface and crustal deformation (Fig. 2, Extended Data Fig. 5). Vertical motions in the crust and lithosphere that overlie the plume change with time. The surface shows a sequence of events as follows: (1) a short (<0.25 Myr) dynamic uplift; (2) a large subsidence, which is first caused by mechanical erosion (thinning) of the LAB by mantle flow above the upwelling plume, and then amplified by gravitational down-warping of the LAB caused by fast Rayleigh–Taylor instability in cold dense mantle lithosphere^{6,12}; (3) a surface uplift (at ~ 0.5 –1.0 Myr) resulting from buoyancy and the dynamic push of the plume; and (4) a subsidence in the middle of the uplifted zone due to continuing interplay between plume and lithosphere mantle. The uplift/subsidence events alternate over several Myr as plume ponds at the LAB. Small initial radial crustal cracks (not distinguishable in Fig. 2b) vanish after ~ 0.25 Myr.

The predictions from the above 'end-member' cases depart strongly from those observed in our next experiment, which considers the mixed scenario of 'active/passive rifting'⁸, in which the lithosphere is submitted to ultra-slow far-field extension and simultaneously impinged upon by an upwelling mantle plume (Figs 3, 4, Extended Data Fig. 6). The initial radial uplift patterns disappear within the first 0.5 Myr, and a series of linear localized sub-parallel normal faults form in the direction normal to the direction of extension. When the plume impacts the base of the lithosphere, the lithosphere begins to thin in a single direction that is

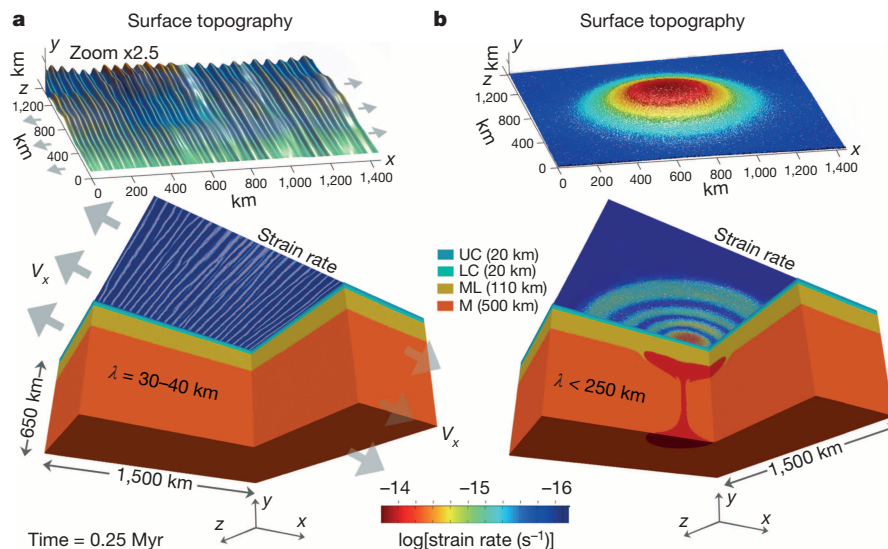


Figure 2 | Initial stages of end-member 3D numerical experiments. Top row, surface topography resulting from end-member cases of plume–lithosphere interactions (bottom row). Shown also in the bottom row are corresponding surface strain rate fields and the material phase field. **a**, No plume, and ultra-slow unidirectional tectonic extension (at a rate $V_x = 3 \text{ mm yr}^{-1}$) applied to a 250-Myr-old continental lithosphere. Note distributed closely spaced sub-parallel small-offset fault patterns spaced at $\lambda \approx 30\text{--}40 \text{ km}$. **b**, A 200-km plume initialized at 650-km depth impinges on the bottom of non-prestressed lithosphere. Note axisymmetric surface undulations with wavelengths (λ) of $<250 \text{ km}$ (150–250 km). Surface strain rate fields, or distributions, characterize the rates of surface deformation and topography evolution, ranging from fast ($>10^{-14} \text{ s}^{-1}$) to slow ($<10^{-16} \text{ s}^{-1}$). ML, mantle lithosphere; UC, upper crust; LC, lower crust; M, sub-lithosphere mantle. See also Extended Data Figs 1, 5, 7. The topography in **a** has a 2.5 vertical exaggeration compared to **b**.

determined by the far-field stress field and brittle localization in the lithosphere. The large-offset faults become increasingly important above the plume head, while almost no faulting occurs at distance larger than 200 km from the centre of rifting. Faults soon merge into a localized, linear rift, which opens at a significantly faster (5 to 10 times) rate than that of the ultra-slow passive extension applied at the borders (a 80-km-wide rift forms in less than 2.5 Myr while the amount of extension applied at the borders is only 7.5 km), showing that in this case the major driving force of rifting is active mantle upwelling. As predicted⁷, rifting creates a rift-parallel conduit for the flow of the plume material away from the region underlain by the starting plume head, further enhancing localized thinning of the lithosphere. The large-scale deformation becomes ‘cylindrical’, that is, two-dimensional, resembling a ‘typical’ continental rift (Extended Data Figs 5b, 6). The evolution of topography is summarized in Fig. 4: the initial surface response ($<1.5 \text{ Myr}$) is axisymmetric, as in the case of Fig. 2b, with large-scale subsidence evolving into a large-scale 1-km-high dome that narrows (300 km) and becomes more elevated with time ($>1.5 \text{ km}$). During the following few Myr, small-scale faulting is superimposed on a long-wavelength topography, which vanishes after 5–6 Myr. Further surface topography evolution is dominated by localized tectonic-scale rift structures. It is noteworthy that pre-rift doming is largely reduced.

In the next series of 3D experiments (Extended Data Fig. 7, 8), far-field velocities were applied at four sides of the model to produce a combination of simple and pure shear. The plume–lithosphere interaction then triggers formation of a sub-diagonal large-scale normal-and-strike-slip fault segmented by oblique Riedel shears. The fault has small lateral displacement ($<10\%$ of its length) and nucleates by the simultaneous coalescence of small-scale shears along the entire length of the rift. Interestingly, the surface topography resembles an early stage of a segmented spreading centre with a strong strike-slip component (for example, the Dead Sea rift and the Gulf of Aden). The same experiment implemented without mantle upwelling (Extended Data Fig. 7) shows only a series of distributed small-scale conjugated fault patterns.

We conclude from our modelling that localization of large-scale linear normal and strike-slip faults can be triggered and maintained by mantle flow that impacts the base of a pre-stressed lithosphere, so that the final state of the rifted lithosphere is an indicator of the far-field stress at the time the plume arrived. This suggests an efficient mechanism for continental rift initiation and breakup that involves passive and active rifting processes that interact with each other⁸, resulting in the development of large continental rifts (for example, the Afar, Gulf of Aden, Dead Sea, Baikal and East African rifts^{26,27}) and plate-scale strike-slip faults (for example, the Karakoum and North Anatolian faults¹⁰).

There is also a significant difference in the impact of the rheological profile on rifting style in the case of dominant active rifting compared to dominant passive rifting. Narrow rifting, conventionally attributed to cold strong lithosphere in passive rifting mode¹⁴, may develop in

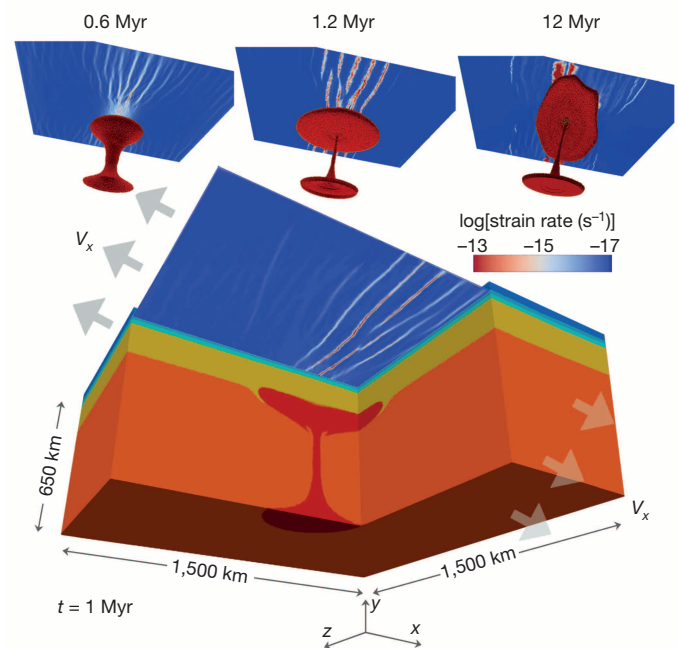


Figure 3 | Surface fault patterns and strain distributions in the case of plume–lithosphere interaction in the presence of a weak tectonic far-field stress field ($V_x = 3 \text{ mm yr}^{-1}$). The bottom panel shows the surface strain rate field (distribution) and the material phase field at time $t = 1 \text{ Myr}$. The top row shows the time evolution of the surface strain rate field (white to red colour, fast surface topography movements) at several key stages ($t = 0.6 \text{ Myr}$, 1.2 Myr , 12 Myr) of plume–lithosphere interaction. See also Fig. 4 and Extended Data Fig. 6. Note progressive focusing and amplification of surface deformation (linear faults) above the plume, resulting in rapid localization of non-axisymmetric, uni-directional deformation (rifting) at the surface. Rifting results in the alignment of the plume with the rift axis, causing rift-parallel flow of plume material away from the region underlain by the starting plume head, which further accelerates localized thinning of the lithosphere and the lithosphere break-up process. The lithosphere deformation soon becomes ‘cylindrical’, almost two-dimensional, resembling a ‘typical’ continental rift.

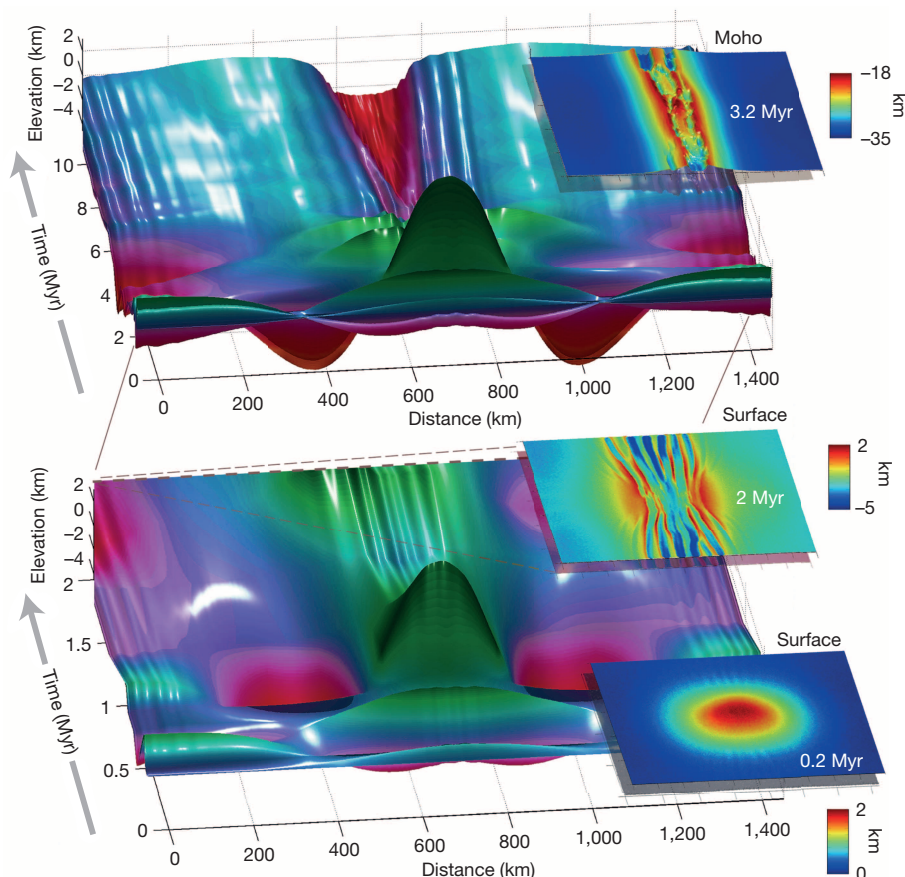


Figure 4 | Summary of time evolution of representative topography profile x, y . Top panel, overall time evolution of surface topography along a profile that runs through the centre of the model in the direction of extension (experiment shown in Fig. 3). Bottom panel, a zoom to the first, most important stages (1–2 Myr) of surface evolution. The rectangular insets show surface and Moho topography at key stages of time evolution (see Extended Data Fig. 6 for more details); colour bars with units of km correspond to vertical amplitudes of these topographies. Note initially short-lived axisymmetric domal uplift, which is rapidly followed by narrow, almost two-dimensional, rifting patterns. The lineation of these rifts is normal to the tectonic stress field direction. After 2 Myr, surface deformation patterns have shorter wavelengths than Moho topography.

weak hot ultra-stretched lithosphere during active rifting, after plume impingement on a tectonically pre-stressed lithosphere. In that case, initially ultra-wide small-amplitude rift patterns focus, in few Myr, into large-scale faults that form a narrow rift. Also, wide rifting may develop during ultra-slow spreading of strong lithosphere, and ‘switch’ to narrow rifting upon plume impingement.

Interplay between mantle–lithosphere interactions, rheological structure and intraplate stresses can result in two main topographic evolution scenarios: (1) in the general case of rheologically stratified lithosphere (with thermo-tectonic ages of 50–500 Myr; ref. 13), the topography exhibits short wavelengths ($\lambda \approx 50, 100, 350$ km) that are much smaller than those of mantle flow ($\lambda > 1,000$ km). In the presence of even a weak pre-existing far-field forcing, the topography is strongly asymmetric and is shaped by faults and rifts. (2) In the case of young and weak (age < 50 Myr) or old and strong plates (> 700 Myr ago, shields), the dynamic topography has ‘classical’ features, that is, it correlates with large-scale mantle flow.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 5 March; accepted 18 July 2014.

1. Lithgow-Bertelloni, C. & Silver, P. G. Dynamic topography, plate driving forces and the African superswell. *Nature* **395**, 269–272 (1998).
2. Moucha, R. & Forte, A. M. Changes in African topography driven by mantle convection. *Nature Geosci.* **4**, 707–712 (2011).
3. Hager, B. H., Clayton, R. W., Richards, M. A., Comer, R. P. & Dziewonski, A. M. Lower mantle heterogeneity, dynamic topography and the geoid. *Nature* **313**, 541–545 (1985).
4. Van der Hilst, R. D., Widiyantoro, S. & Engdahl, E. R. Evidence for deep mantle circulation from global tomography. *Nature* **386**, 578–584 (1997).
5. Guillou-Frotier, L., Burov, E., Nehlig, P. & Wyns, R. Deciphering plume-lithosphere interactions beneath Europe with topographic signatures. *Glob. Planet. Change* **58**(Spec. vol. on topography of Europe), 119–140 (2007).

6. Burov, E. & Guillou-Frotier, L. The plume head-lithosphere interaction using a tectonically realistic formulation for the lithosphere. *Geophys. J. Int.* **161**, 469–490 (2005).
7. Sleep, N. Lateral flow and ponding of starting plume material. *J. Geophys. Res.* **102**, 10001–10012 (1997).
8. Courtillot, V., Jaupart, C., Manighetti, I., Tapponnier, P. & Besse, J. On causal links between flood basalts and continental breakup. *Earth Planet. Sci. Lett.* **166**, 177–195 (1999).
9. Flament, N., Gurnis, M. & Müller, R. D. A review of observations and models of dynamic topography. *Lithosphere* **5**, 189–210 (2013).
10. Faccenna, C. & Becker, T. W. Shaping mobile belts by small-scale convection. *Nature* **465**, 602–605 (2010).
11. Anderson, D. L. The sublithospheric mantle as the source of continental flood basalt: the case against the continental lithosphere and plume head reservoirs. *Earth Planet. Sci. Lett.* **123**, 269–280 (1994).
12. Molnar, P. & Houseman, G. A. The effects of buoyant crust on the gravitational instability of thickened mantle lithosphere at zones of intracontinental convergence. *Geophys. J. Int.* **158**, 1134–1150 (2004).
13. Burov, E. B. Rheology and strength of the lithosphere. *Mar. Petrol. Geol.* **28**, 1402–1443 (2011).
14. Buck, R. in *Treatise on Geophysics* Vol. 6, *Crust and Lithosphere Dynamics* (ed. Watts, T.) 335–376 (Elsevier, 2007).
15. Burov, E., Guillou-Frotier, L., d’Acremont, E., Le Pourhiet, L. & Cloetingh, S. Plume head–lithosphere interactions near intra-continental plate boundaries. *Tectonophysics* **434**, 15–38 (2007).
16. Sleep, N. H., Ebinger, C. J. & Kendall, J.-M. in *The Early Earth: Physical, Chemical and Biological Development* (eds Fowler, C. M. R., Ebinger, C. J. & Hawkesworth, C. J.) 135–150 (Geological Society of London Spec. Publ. 199, 2002).
17. Lin, S. C. & van Keken, P. E. Multiple volcanic episodes of flood basalts caused by thermochemical mantle plumes. *Nature* **436**, 250–252 (2005).
18. Boher, M., Abouchami, W., Michard, A., Albarède, F. & Arndt, N. T. Crustal growth in west Africa at 2.1 Ga. *J. Geophys. Res.* **97**, 345–369 (1992).
19. Lustrino, M. & Carminati, E. Phantom plumes in Europe and the circum-Mediterranean region. *Geol. Soc. Am. Spec. Publ.* **430**, 723–745 (2007).
20. Molnar, P. & Houseman, G. A. Rayleigh–Taylor instability, lithospheric dynamics, surface topography at convergent mountain belts, and gravity anomalies. *J. Geophys. Res.* **118**, 2544–2557 (2013).
21. Foulger, G. R. *et al.* The seismic anomaly beneath Iceland extends down to the mantle transition zone and no deeper. *Geophys. J. Int.* **142**, F1–F5 (2000).
22. Sun, D., Helmberger, D. & Gurnis, M. A narrow mid-mantle plume below southern Africa. *Geophys. Res. Lett.* **37**, L09302 (2010).

23. Collins, J. A., Wolfe, C. J. & Laske, G. Shear wave splitting at the Hawaiian hot spot from the PLUME land and ocean bottom seismometer deployments. *Geochem. Geophys. Geosyst.* **13**, Q02007 (2012).
 24. Davies, G. F. Temporal variation of the Hawaiian plume flux. *Earth Planet. Sci. Lett.* **113**, 277–286 (1992).
 25. Wessel, P. & Keating, B. H. Temporal variations of flexural deformation in Hawaii. *J. Geophys. Res.* **99**, 2747–2756 (1994).
 26. Pik, R., Marty, R. & Hilton, D. R. How many mantle plumes in Africa? The geochemical point of view. *Chem. Geol.* **226**, 100–114 (2006).
 27. Moreira, M., Valbracht, P. J., Staudacher, T. & Allègre, C. J. Rare gas systematics in Red Sea ridge basalts. *Geophys. Res. Lett.* **23**, 2453–2456 (1996).
 28. Ueda, K., Gerya, T. & Sobolev, S. V. Subduction initiation by thermal–chemical plumes: numerical studies. *Phys. Earth Planet. Inter.* **171**, 296–312 (2008).
 29. Burov, E. & Cloetingh, S. Controls of mantle plumes and lithospheric folding on modes of intra-plate continental tectonics: differences and similarities. *Geophys. J. Int.* **37**, L03309 (2010).
 30. Gerya, T. V. *Introduction to Numerical Geodynamic Modelling* (Cambridge Univ. Press, 2010).
- Acknowledgements** B. Evans and P. Molnar are thanked for discussions. A. B. Watts is thanked for comments on the manuscript and corrections to it. This study was co-funded by an Advanced ERC grant RHEOLITH (E.B.), by INSU-CNRS, by a UPMC Invited Professor grant (T.G.) and by an ETH Invited Professor grant (E.B.). Numerical simulations were performed on the ETH Brutus cluster and on the ERC-funded SGI Ulysse cluster of ISTEP (UPMC). Open source software ParaView (<http://www.paraview.org>) was used for 3D visualization.
- Author Contributions** E.B. designed the study, conducted some of the 3D experiments, designed the 2D thermo-mechanical code and conducted 2D experiments. T.G. designed the 3D thermo-mechanical code and conducted some of the 3D experiments. Both authors discussed problems and methods, interpreted the data and wrote the paper.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.B. (evgenii.burov@upmc.fr).

RIPK1 maintains epithelial homeostasis by inhibiting apoptosis and necroptosis

Marius Dannappel^{1*}, Katerina Vlantis^{1*}, Snehlata Kumari^{1*}, Apostolos Polykratis^{1*}, Chun Kim¹, Laurens Wachsmuth¹, Christina Eftychi¹, Juan Lin¹, Teresa Corona¹, Nicole Hermance², Matija Zelic², Petra Kirsch³, Marijana Basic⁴, Andre Bleich⁴, Michelle Kelliher^{2§} & Manolis Pasparakis^{1§}

Necroptosis has emerged as an important pathway of programmed cell death in embryonic development, tissue homeostasis, immunity and inflammation^{1–8}. RIPK1 is implicated in inflammatory and cell death signalling^{9–13} and its kinase activity is believed to drive RIPK3-mediated necroptosis^{14,15}. Here we show that kinase-independent scaffolding RIPK1 functions regulate homeostasis and prevent inflammation in barrier tissues by inhibiting epithelial cell apoptosis and necroptosis. Intestinal epithelial cell (IEC)-specific RIPK1 knockout caused IEC apoptosis, villus atrophy, loss of goblet and Paneth cells and premature death in mice. This pathology developed independently of the microbiota and of MyD88 signalling but was partly rescued by TNFR1 (also known as TNFRSF1A) deficiency. Epithelial FADD ablation inhibited IEC apoptosis and prevented the premature death of mice with IEC-specific RIPK1 knockout. However, mice lacking both RIPK1 and FADD in IECs displayed RIPK3-dependent IEC necroptosis, Paneth cell loss and focal erosive inflammatory lesions in the colon. Moreover, a RIPK1 kinase inactive knock-in delayed but did not prevent inflammation caused by FADD deficiency in IECs or keratinocytes, showing that RIPK3-dependent necroptosis of FADD-deficient epithelial cells only partly requires RIPK1 kinase activity. Epidermis-specific RIPK1 knockout triggered keratinocyte apoptosis and necroptosis and caused severe skin inflammation that was prevented by RIPK3 but not FADD deficiency. These findings revealed that RIPK1 inhibits RIPK3-mediated necroptosis in keratinocytes *in vivo* and identified necroptosis as a more potent trigger of inflammation compared with apoptosis. Therefore, RIPK1 is a master regulator of epithelial cell survival, homeostasis and inflammation in the intestine and the skin.

Mice lacking RIPK1 die perinatally, exhibiting apoptosis in multiple tissues including the intestine⁹ (Extended Data Fig. 1a). We generated mice with IEC-specific knockout of RIPK1 (RIPK1^{IEC-KO}) (Fig. 1a and Extended Data Fig. 1b–d) and found that they showed reduced body weight and died within the first 4 weeks of life (Fig. 1b, c). Intestinal sections from 3-week-old RIPK1^{IEC-KO} mice revealed pronounced villus atrophy and Paneth cell loss in the ileum, as well as reduced numbers of goblet cells and differentiated enterocytes, crypt elongation and IEC hyperproliferation in both the ileum and colon (Fig. 1d and Extended Data Fig. 1e, f). RIPK1^{IEC-KO} intestines contained highly increased numbers of cleaved caspase-3-positive (CC3⁺) apoptotic IECs, but did not show signs of epithelial erosion, suggesting that the epithelial barrier remained largely intact (Fig. 1d and Extended Data Fig. 1e, f). Messenger RNA levels of *Tnf*, *Il1b*, *Ccl5* and *Cxcl1* and leukocyte numbers were moderately increased in the ileum and colon of RIPK1^{IEC-KO} mice (Fig. 1e, f and Extended Data Fig. 2), indicating the presence of a mild inflammatory response. Newborn RIPK1^{IEC-KO} mice showed only a few apoptotic IECs, while 7-day-old mice showed increased numbers of apoptotic IECs, loss of goblet cells, elongated hyperproliferative crypts

and increased immune cell infiltration concomitant with moderately increased expression of tumour necrosis factor (TNF) and interleukin (IL)-1 β (Extended Data Fig. 3). Tamoxifen-inducible ablation of RIPK1 in IECs of adult mice (RIPK1^{tamIEC-KO}) caused rapid weight loss and death of the animals (Extended Data Fig. 4a–d) due to extensive IEC apoptosis resulting in severe intestinal pathology (Extended Data Fig. 4e–i). Collectively, these results revealed that epithelial-cell-intrinsic RIPK1 is essential for IEC survival and the maintenance of intestinal tissue structure and homeostasis.

The luminal microbiota regulates intestinal homeostasis and contributes to inflammation both in human inflammatory bowel disease and in mouse models^{16,17}. Antibiotic treatment briefly delayed but did not prevent severe intestinal pathology, weight loss and death of mice with inducible or constitutive IEC-specific RIPK1 knockout (Extended Data Fig. 4c–e, j–l). Moreover, RIPK1^{IEC-KO} mice raised under germ-free conditions developed similar intestinal pathology, although some of the mice survived up to the age of 5 weeks, indicating that the microbiota contributes to disease severity (Fig. 1g–j). In addition, MyD88 deficiency did not prevent intestinal pathology in RIPK1^{IEC-KO} mice (Fig. 2a, b, d and Extended Data Fig. 5a–c). Therefore, epithelial cell death and intestinal pathology in RIPK1^{IEC-KO} mice develop independently of the microbiota and of MyD88-dependent TLR signalling. RIPK1^{IEC-KO}/Tnfr1^{−/−} mice were partially protected from severe wasting, with approximately 50% of the mice surviving to at least 3 months of age, and showed reduced IEC apoptosis and ameliorated intestinal pathology (Fig. 2a–e and Extended Data Fig. 5a–c). Thus, IEC death and intestinal pathology in RIPK1^{IEC-KO} mice is induced mainly by TNFR1 signalling but TNFR1-independent pathways also contribute.

Epithelial FADD deficiency inhibited IEC apoptosis and villus atrophy and prevented the severe wasting and premature death of RIPK1^{IEC-KO} mice (Fig. 2f–h), suggesting an important pathogenic role of FADD-dependent apoptosis. However, RIPK1^{IEC-KO}/FADD^{IEC-KO} intestines contained increased numbers of CC3⁺ dying IECs and showed Paneth cell loss in the ileum and signs of colitis affecting primarily the distal colon (Fig. 2f, i, j and Extended Data Fig. 6). In contrast with RIPK1^{IEC-KO} mice, in which the distal colon was only mildly affected and that did not show epithelial erosions (0 of 18 mice analysed), histological examination of Swiss roll sections revealed focal epithelial erosions in the distal colon of RIPK1^{IEC-KO}/FADD^{IEC-KO} mice (8 of 14 mice) and RIPK1^{IEC-KO} mice with heterozygous epithelial FADD deficiency (FADD^{hetIEC-KO}) (5 of 6 mice) at the age of 3 weeks (Extended Data Fig. 6). Adult RIPK1^{IEC-KO}/FADD^{IEC-KO} mice displayed villus shortening, loss of goblet and Paneth cells and crypt elongation in the small intestine, as well as focal inflammatory erosive lesions and crypt abscess formation in the distal colon in 5 of 7 mice examined (Extended Data Fig. 7a). Therefore, although epithelial FADD deficiency strongly reduced the number of apoptotic

¹Institute for Genetics, Centre for Molecular Medicine (CMMC), and Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, 50931 Cologne, Germany. ²Department of Cancer Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA. ³Tierforschungszentrum, University of Ulm, Albert-Einstein-Allee 11, D-89081 Ulm, Germany. ⁴Institute for Laboratory Animal Science, Hannover Medical School, D-30625 Hannover, Germany.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

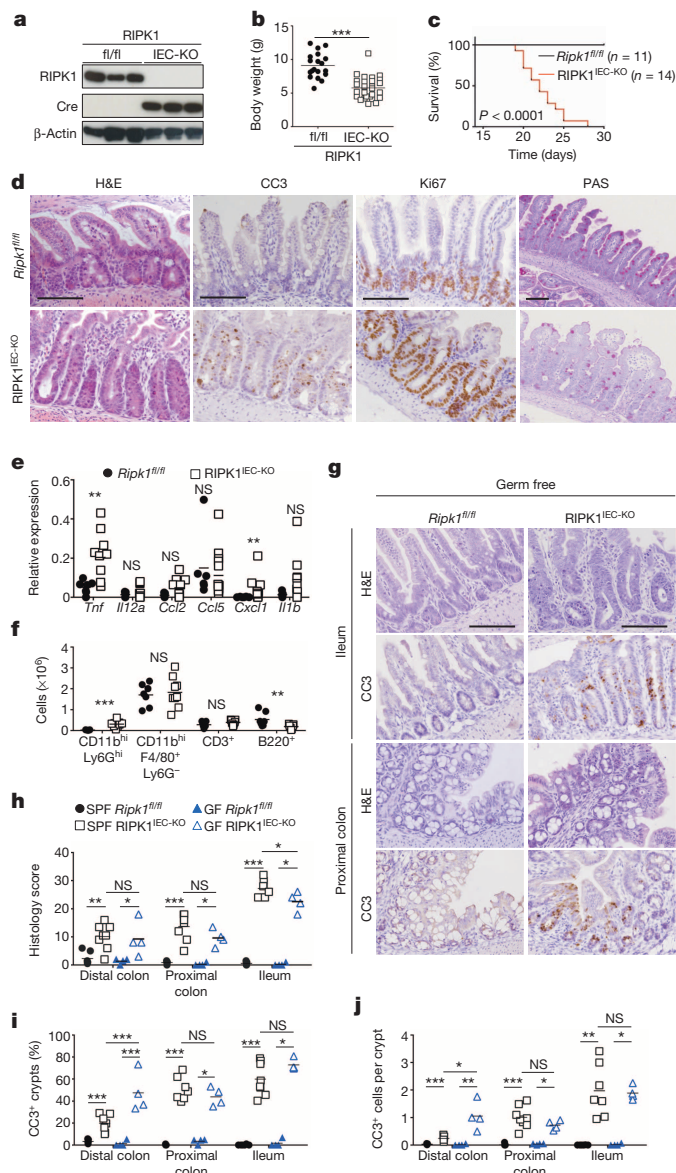


Figure 1 | Epithelial RIPK1 ablation causes microbiota-independent intestinal pathology. **a**, Immunoblot of small intestine IECs from *Ripk1^{fl/fl}* (fl/fl) and *RIPK1^{IEC-KO}* (IEC-KO) mice. **b**, **c**, Body weight (**b**) and Kaplan-Meier survival curve (**c**) of *Ripk1^{fl/fl}* and *RIPK1^{IEC-KO}* mice. **d**, Representative images of ileal sections from *Ripk1^{fl/fl}* and *RIPK1^{IEC-KO}* mice stained as indicated. H&E, haematoxylin and eosin. **e**, **f**, Quantitative polymerase chain reaction with reverse transcription (qRT-PCR) of cytokine and chemokine expression (**e**) and fluorescence-activated cell sorting (FACS) of lamina propria leukocytes (**f**) in the small intestine of *Ripk1^{fl/fl}* and *RIPK1^{IEC-KO}* mice. **g**, Representative images of H&E- or CC3-stained intestinal sections from germ-free *Ripk1^{fl/fl}* and *RIPK1^{IEC-KO}* mice. **h**–**j**, Quantification of histological pathology score (**h**), crypts containing CC3⁺ cells (**i**) and CC3⁺ cells per crypt (**j**) in intestinal sections of the indicated mice. GF, germ free; SPF, specific pathogen free. Scale bars, 100 μ m. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.005$; NS, not significant.

epithelial cells and prevented wasting and premature death of *RIPK1^{IEC-KO}* mice, it did not normalize but rather altered their intestinal pathology, with *RIPK1^{IEC-KO}/FADD^{IEC-KO}* animals resembling the phenotype of *FADD^{IEC-KO}* mice⁵.

We reasoned that RIPK3-dependent necroptosis could induce the death of IECs in *RIPK1^{IEC-KO}/FADD^{IEC-KO}* mice, as in *FADD^{IEC-KO}* mice⁵. Indeed, triple-deficient *RIPK1^{IEC-KO}/FADD^{IEC-KO}/Ripk3^{-/-}* mice did not show macroscopic or histological signs of intestinal pathology, although they displayed slightly increased mRNA levels of *Tnf*, *Il1b*, *Ccl5*

and *Il10* in the colon (Fig. 2f–k and Extended Data Figs 5d, e, 6, 7a, b). In contrast, RIPK3 deficiency alone did not prevent wasting, early lethality, IEC apoptosis or intestinal pathology in *RIPK1^{IEC-KO}* mice (Fig. 2f–j and Extended Data Figs 5d, e, 6, 7b). Therefore, RIPK3-dependent necroptosis does not constitute a primary pathway inducing IEC death and intestinal pathology in *RIPK1^{IEC-KO}* mice, but becomes functionally important when FADD ablation inhibits apoptosis and sensitizes RIPK1-deficient IECs to necroptosis. These results suggested that RIPK3 induces RIPK1-independent necroptosis in FADD-deficient IECs, which is surprising as RIPK1 kinase activity is considered essential for RIPK3-mediated necroptosis, at least downstream of TNFR1 (refs 14, 15), which is the major driver of epithelial cell necroptosis in the colon of *FADD^{IEC-KO}* mice⁵. To address specifically the role of RIPK1 kinase activity, we crossed *FADD^{IEC-KO}* mice with knock-in mice expressing kinase-inactive RIPK1 (*RIPK1* (D138N)) and found that lack of RIPK1 kinase activity did not prevent IEC necroptosis and intestinal inflammation in *FADD^{IEC-KO}* mice (Fig. 2l and Extended Data Fig. 7c, d). Therefore, RIPK1 deficiency or loss of its kinase activity only partly inhibited RIPK3-dependent necroptosis of FADD-deficient IECs, providing genetic evidence that both RIPK1-dependent and RIPK1-independent pathways drive RIPK3-mediated necroptosis and intestinal inflammation in *FADD^{IEC-KO}* mice.

As shown previously^{9,18}, RIPK1 deficiency but not loss of its kinase activity partially inhibited TNF-mediated NF- κ B activation in mouse embryonic fibroblasts (MEFs) (Extended Data Fig. 8a), suggesting that loss of RIPK1 scaffolding function might trigger epithelial cell apoptosis by inhibiting NF- κ B. However, NF- κ B inhibition in the intestinal epithelium did not phenocopy the intestinal pathology of *RIPK1^{IEC-KO}* mice^{19–21}. Moreover, sustained NF- κ B activation in RIPK1-deficient epithelial cells by expression of a constitutively active *Ikk2* transgene^{22,23} did not prevent IEC death and rather aggravated the intestinal pathology of *RIPK1^{IEC-KO}* mice, suggesting that RIPK1 deficiency does not sensitize IECs to apoptosis by inhibiting NF- κ B (Extended Data Fig. 8b–d).

RIPK1-deficient IECs did not show reduced mRNA expression of pro-survival genes, but expressed reduced amounts of cIAP1 (also known as BIRC2) and TRAF2 proteins (Fig. 3a, b), suggesting that RIPK1 might control their stability. Indeed, consistent with previous reports^{24,25}, we found that TNF stimulation induced rapid degradation of TRADD, cIAP1, TRAF2 and c-FLIP (also known as CFLAR), but not of TRAF6, in RIPK1-deficient but not in *RIPK1* (D138N) primary MEFs (Fig. 3c), suggesting that lack of the RIPK1 scaffolding function triggers degradation of proteins recruited to the TNFR1 signalling complex. We then sought to address whether RIPK1-deficient intestinal epithelial cells show similar responses. However, small intestinal crypts from *RIPK1^{IEC-KO}* mice died shortly after isolation and failed to form organoids (data not shown). Nevertheless, tamoxifen-inducible RIPK1 deletion resulted in reduced expression of cIAP1, TRAF2 and c-FLIP proteins and caused rapid organoid death (Fig. 3d–f), suggesting that loss of pro-survival protein expression contributes to death of RIPK1-deficient IECs. Interestingly, TRAF2-deficient mice develop a similar intestinal phenotype and early lethality that was not prevented by antibiotic treatment and was only partially rescued by TNFR1 deficiency, suggesting that loss of TRAF2 could contribute to the intestinal pathology of *RIPK1^{IEC-KO}* mice²⁶.

SMAC (also known as DIABLO) mimetic-compound-mediated cIAP1/2 degradation induced activation of both non-canonical and canonical NF- κ B and the expression of TNF and other cytokines^{27,28}, suggesting that loss of cIAP proteins could trigger the expression of TNF and other cytokines by RIPK1-deficient IECs that could act in an autocrine fashion to induce their death. Indeed, RIPK1-deficient primary IECs showed increased NF- κ B activation and expressed elevated levels of TNF, while tamoxifen-inducible deletion of RIPK1 upregulated TNF expression in organoids (Fig. 3g–j). Moreover, intestinal crypts from *RIPK1^{IEC-KO}/Tnfr1^{-/-}* mice did not die rapidly, as did crypts from *RIPK1^{IEC-KO}* mice, and although they did not grow as well as control organoids they could be maintained for a short period in culture (Fig. 3k), consistent with a role of autocrine TNF production in triggering the death of RIPK1-deficient IECs. RIPK1/TNFR1 double-deficient organoids were more sensitive to

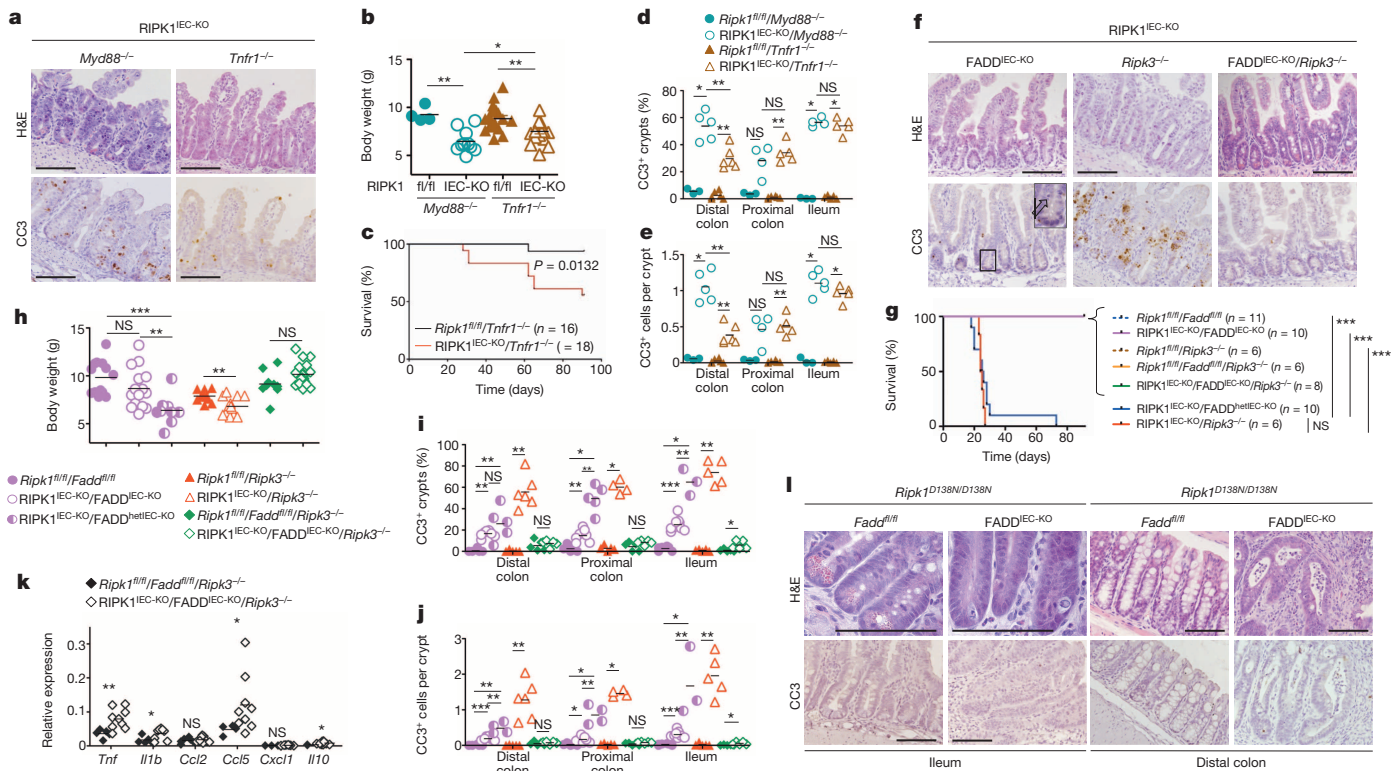


Figure 2 | Death of RIPK1-deficient IECs depends on FADD and RIPK3. **a**, Representative images of H&E- or CC3-stained ileal sections from the indicated mice. **b, c**, Body weight (b) and Kaplan–Meier survival curve (c) of the indicated mice. **d, e**, Quantification of crypts containing CC3⁺ cells (d) and CC3⁺ cell numbers per crypt (e) in intestinal sections of the indicated mice. **f**, Representative images of ileal sections from the indicated mice. Arrow depicts CC3⁺ dying IEC. **g, h**, Kaplan–Meier survival curve (g) and body weight (h) of

death induced by polyinosinic:polycytidylic acid (poly(I:C)), interferon (IFN)- β and IFN- γ (Fig. 3k), indicating that TIR-domain-containing adapter-inducing interferon- β (TRIF)- and IFN-dependent pathways also contribute to the death of RIPK1-deficient IECs. However, IEC-specific ablation of TRIF did not considerably ameliorate intestinal pathology in RIPK1^{IEC-KO} mice (Extended Data Fig. 8e–h), suggesting that TRIF-mediated signalling is not a major driver of IEC apoptosis but could contribute in a redundant fashion together with TNFR1 and IFN signalling. Together, these findings suggest that RIPK1 deficiency results in degradation of cIAP proteins and TRAF2, triggering NF- κ B-mediated expression of TNF and other cytokines that act in a partly redundant fashion to induce the death of RIPK1-deficient IECs. The mild upregulation of *Tnf* in the colon of RIPK1^{IEC-KO}/FADD^{IEC-KO}/Ripk3^{-/-} mice (Fig. 2k) is compatible with this epithelial cell death independent mechanism inducing *Tnf* expression, which also provides a microbiota-independent trigger of IEC death in RIPK1^{IEC-KO} mice.

To investigate the function of RIPK1 in the epidermis, we generated mice with keratinocyte-specific RIPK1 deficiency (RIPK1^{E-KO} mice) (Extended Data Fig. 9a). RIPK1^{E-KO} mice progressively developed severe inflammatory skin lesions starting at around 1 week of age, characterized by epidermal thickening, increased keratin 14 (K14; also known as KRT14) and reduced K10 expression and upregulation of K6 in the interfollicular epidermis, as well as infiltration of Gr-1-positive cells and increased inflammatory cytokine and chemokine expression (Fig. 4a–c and Extended Data Fig. 9a, b). Histological analysis revealed the presence of CC3⁺ but also CC3⁻ dying keratinocytes, identified by their pyknotic nuclei and eosinophilic cytoplasm in the epidermis of RIPK1^{E-KO} mice (Fig. 4d and Extended Data Fig. 9d). To address the role of FADD-mediated apoptosis, we generated FADD^{E-KO}/RIPK1^{E-KO} mice and found that they did not show early postnatal skin inflammation and death, as did FADD^{E-KO}

mice with the indicated genotypes. **i, j**, Quantification of crypts containing CC3⁺ cells (i) and CC3⁺ cell numbers per crypt (j) in intestinal sections of the indicated mice. **k**, qRT–PCR analysis of cytokine and chemokine expression in the small intestine of the indicated mice. **l**, Representative images of intestinal sections from 4-month-old mice with the indicated genotypes. Scale bars, 100 μ m. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.005$; NS, not significant.

mice⁴, but developed severe inflammatory skin lesions later, similarly to RIPK1^{E-KO} mice (Fig. 4a–c and Extended Data Fig. 9c, d). Importantly, FADD^{E-KO}/Ripk1^{D138N/D138N} mice showed a similar skin phenotype to FADD^{E-KO}/RIPK1^{E-KO} mice (Fig. 4a, b). Therefore, lack of RIPK1 or its kinase activity delayed but could not prevent keratinocyte necroptosis and inflammation caused by FADD deficiency. Additional deficiency in RIPK3 fully prevented skin inflammation in RIPK1^{E-KO}/FADD^{E-KO}/Ripk3^{-/-} mice (Fig. 4a–c and Extended Data Fig. 9e, g), showing that RIPK1-independent RIPK3-dependent keratinocyte necroptosis drives skin inflammation.

Surprisingly, RIPK1^{E-KO}/Ripk3^{-/-} mice did not develop skin inflammation and their epidermis contained strongly reduced numbers of CC3⁻ but similar numbers of CC3⁺ keratinocytes compared with RIPK1^{E-KO} mice (Fig. 4a–d and Extended Data Fig. 9d, e, g), showing that RIPK3-mediated necroptosis of RIPK1-deficient keratinocytes causes the inflammatory skin lesions. CRISPR/Cas9-mediated knockout of MLKL also prevented skin inflammation in RIPK1^{E-KO} mice, further supporting an essential role of RIPK3/MLKL-dependent necroptosis in this model (Extended Data Fig. 10). Therefore, at least in the context of keratinocyte-specific RIPK1 deficiency, necroptosis of keratinocytes triggers skin inflammation but apoptosis does not, highlighting the different immune regulatory properties of these programmed cell death pathways in a disease-relevant *in vivo* experimental setting. Considering that *Ripk1*^{D138N/D138N} mice did not develop skin lesions (Fig. 4a), these results show that RIPK1 acts through a kinase-independent scaffolding function to restrain RIPK3-mediated necroptosis in keratinocytes and prevent skin inflammation.

RIPK1^{E-KO}/Tnfr1^{-/-} mice showed largely normal skin until postnatal day (P)28, but developed patchy inflammatory skin lesions by the age of 7–8 weeks (Fig. 4a–d and Extended Data Fig. 9d, e, g). Therefore, TNFR1 critically contributes but is not the only trigger of keratinocyte necroptosis

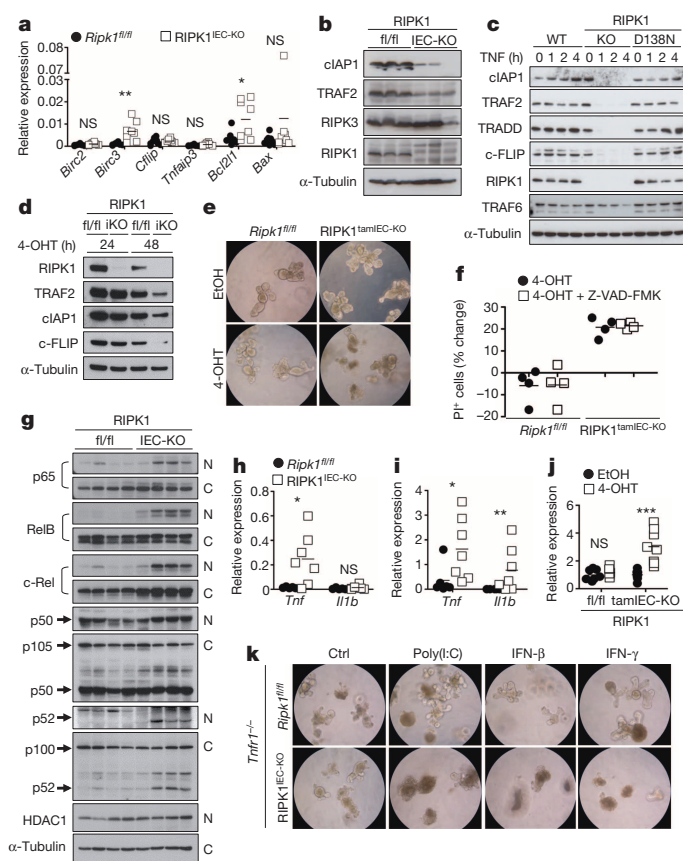


Figure 3 | RIPK1 prevents the degradation of pro-survival proteins.

a, b, qRT-PCR of pro-survival genes (**a**) and immunoblot of the indicated proteins (**b**) in small intestine IECs from *Ripk1*^{fl/fl} (fl/fl) and *RIPK1*^{IEC-KO} (IEC-KO) mice. **c**, Immunoblot of wild-type (WT), *Ripk1*^{-/-} (KO) and *Ripk1*^{D138N/D138N} (D138N) MEFs stimulated with 10 ng ml⁻¹ recombinant murine TNF. **d**, Immunoblot of small intestine organoids from *Ripk1*^{fl/fl} (fl/fl) and *RIPK1*^{tamIEC-KO} (iKO) mice treated with 4-hydroxytamoxifen (4-OHT) for 24 or 48 h. **e**, *Ripk1*^{fl/fl} and *RIPK1*^{tamIEC-KO} organoids were treated with vehicle (ethanol; EtOH) or 4-OHT for 20 h and photographed 48 h later. Original magnification, ×400. **f**, FACS quantification of propidium iodide (PI)⁺ cells in organoids 43 h after treatment with 4-OHT in the presence or absence of z-VAD-FMK. **g**, Immunoblot of cytoplasmic (C) and nuclear (N) proteins from small intestine IECs of four *Ripk1*^{fl/fl} and four *RIPK1*^{IEC-KO} mice with the indicated antibodies. **h, i**, qRT-PCR of *Tnf* and *Il1b* expression in IECs from small intestine (**h**) or colon (**i**) of *Ripk1*^{fl/fl} and *RIPK1*^{IEC-KO} mice. **j**, qRT-PCR of *Tnf* expression in small intestine organoids from *Ripk1*^{fl/fl} and *RIPK1*^{IEC-KO} mice 30 h after treatment with EtOH or 4-OHT. **k**, Representative images of *Ripk1*^{fl/fl}/*Tnfr1*^{-/-} and *RIPK1*^{IEC-KO}/*Tnfr1*^{-/-} organoids 16 h after treatment with 100 μg ml⁻¹ poly(I:C), 1,000 U ml⁻¹ IFN-β and 1,000 U ml⁻¹ IFN-γ. Original magnification, ×400. Ctrl, control. Representative data are shown of four (**e, f, k**), three (**c**) or two (**d, j**) independent experiments. **P* ≤ 0.05, ***P* ≤ 0.01, ****P* ≤ 0.005; NS, not significant.

and inflammation in *RIPK1*^{E-KO} mice. Keratinocyte-specific TRIF ablation also mildly ameliorated skin lesion development in *RIPK1*^{E-KO} mice (Fig. 4a, b and Extended Data Fig. 9e). Collectively, these results showed that TNFR1- and TRIF-mediated signalling contribute to RIPK3-dependent keratinocyte necroptosis and skin inflammation in *RIPK1*^{E-KO} mice.

cIAP1 was expressed at very low levels at steady state and was undetectable after TNF stimulation in *RIPK1*-deficient primary keratinocytes (Fig. 4e), similar to our findings in IECs. Moreover, *RIPK1*-deficient primary keratinocytes were highly sensitive to TNF-induced death that was inhibited by the caspase inhibitor z-VAD-FMK, demonstrating that *in vitro* these cells died primarily by apoptosis and were resistant to necroptosis even though they expressed RIPK3 (Fig. 4f, g). This result

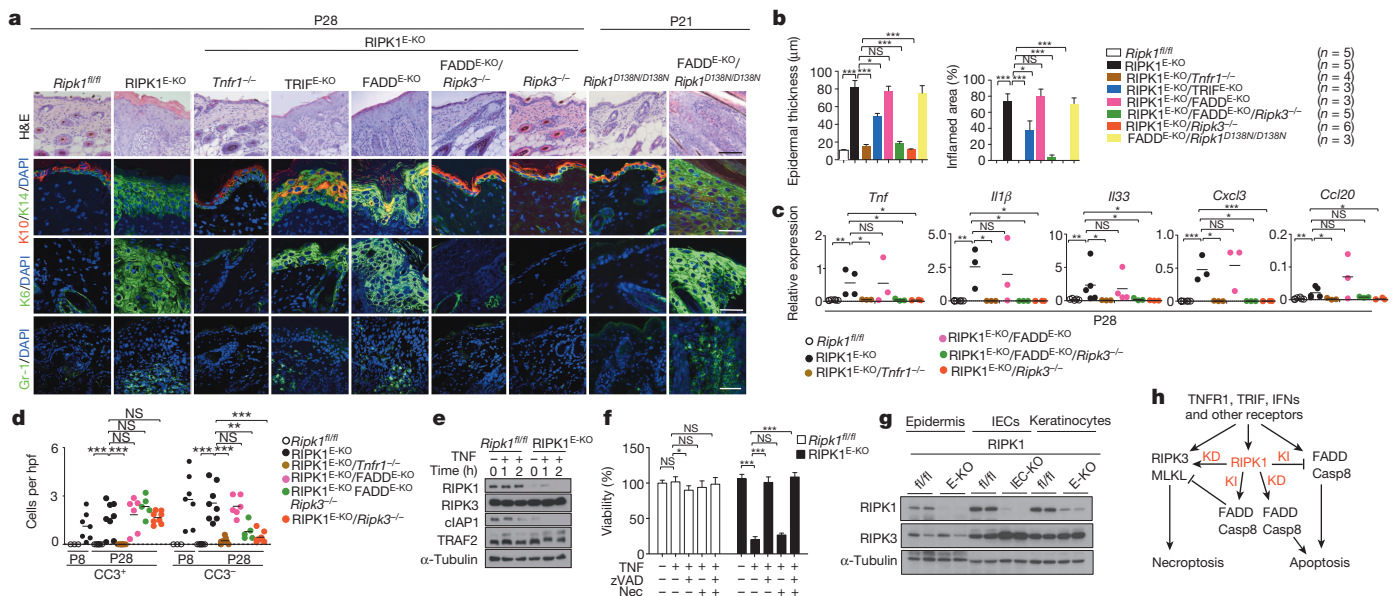


Figure 4 | RIPK1 ablation causes keratinocyte necroptosis and skin inflammation. **a**, Representative images of skin sections from the indicated mice stained with H&E or the indicated antibodies. Nuclei stained with 4',6-diamidino-2-phenylindole (DAPI). Scale bars, 100 μm (H&E); 50 μm (immunostainings). **b**, Microscopic quantification of epidermal thickness and inflamed skin area in the indicated mice. Error bars represent mean values ± standard error of the mean (s.e.m.). **c**, qRT-PCR of cytokine and chemokine expression in total skin from the indicated mice. **d**, Quantification of CC3⁺ and CC3⁻ dying cells per high power field (hpf) in skin sections from the indicated mice. **e**, Immunoblot of primary keratinocytes stimulated with

20 ng ml⁻¹ recombinant murine TNF for the indicated time points. **f**, Quantification of cell viability in primary keratinocytes treated with 50 ng ml⁻¹ recombinant murine TNF in the presence or absence of z-VAD-FMK (zVAD; 20 μM) and necrostatin-1 (Nec; 30 μM). Mean values ± s.e.m. from biological triplicates (*n* = 3) are shown. **g**, Immunoblot of protein extracts from P4 epidermis, IECs or keratinocytes from mice of the indicated genotypes. **h**, Schematic model of the kinase-dependent (KD) and -independent (KI) functions of RIPK1 in apoptosis and necroptosis. Representative data of three independent experiments are shown in **e, f**. **P* ≤ 0.05, ***P* ≤ 0.01, ****P* ≤ 0.005; NS, not significant.

is surprising considering that RIPK1 deficiency *in vivo* triggers skin inflammation by sensitizing keratinocytes to RIPK3-dependent necroptosis, and suggests that mechanisms related to keratinocyte differentiation and the *in vivo* tissue context could determine the sensitivity of keratinocytes to apoptosis and necroptosis.

Collectively, our results revealed a novel kinase-independent scaffolding function of RIPK1 as an inhibitor of epithelial cell apoptosis and necroptosis *in vivo* that is essential for the maintenance of physiological tissue homeostasis in the intestine and the skin. The mechanisms determining whether RIPK1 deficiency sensitizes epithelial cells to apoptosis or necroptosis remain unclear and do not seem to depend on differential expression of RIPK3 (Fig. 4g). Our experiments also showed that necroptosis of FADD-deficient IECs and keratinocytes only partly depends on RIPK1 kinase activity, demonstrating that both RIPK1-dependent and -independent pathways drive RIPK3-mediated necroptosis. Therefore, RIPK1 is a central regulator of cell death that has the capacity to either inhibit or induce apoptosis and necroptosis (Fig. 4h). Considering its key role as a molecular hub integrating and decoding multiple ubiquitination (K63, K48, linear) signals to determine cellular responses to stimuli that are derived (for example, lipopolysaccharide, double-stranded RNA) or induced (for example, TNF, IFNs) by microbes and are capable of triggering both inflammation and cell death, RIPK1 emerges as a key regulator of immunity and homeostasis in barrier tissues.

Note added in proof: An accompanying paper (Takahashi, N. *et al.*, *Nature* <http://dx.doi.org/10.1038/nature13706>) reports similar results on the role of RIPK1 in preventing IEC apoptosis, with some differences on the effect of antibiotic treatment and MYD88 deficiency on the intestinal pathology of RIPK1^{IEC-KO} mice.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 December 2013; accepted 25 June 2014.

Published online 17 August 2014.

- Oberst, A. *et al.* Catalytic activity of the caspase-8-FLIP_L complex inhibits RIPK3-dependent necrosis. *Nature* **471**, 363–367 (2011).
- Kaiser, W. J. *et al.* RIP3 mediates the embryonic lethality of caspase-8-deficient mice. *Nature* **471**, 368–372 (2011).
- Zhang, H. *et al.* Functional complementation between FADD and RIP1 in embryos and lymphocytes. *Nature* **471**, 373–376 (2011).
- Bonnet, M. C. *et al.* The adaptor protein FADD protects epidermal keratinocytes from necroptosis *in vivo* and prevents skin inflammation. *Immunity* **35**, 572–582 (2011).
- Welz, P. S. *et al.* FADD prevents RIP3-mediated epithelial cell necrosis and chronic intestinal inflammation. *Nature* **477**, 330–334 (2011).
- Dillon, C. P. *et al.* Survival function of the FADD-CASPASE-8-cFLIP(L) complex. *Cell Rep.* **1**, 401–407 (2012).
- Duprez, L. *et al.* RIP kinase-dependent necrosis drives lethal systemic inflammatory response syndrome. *Immunity* **35**, 908–918 (2011).
- Upton, J. W., Kaiser, W. J. & Mocarski, E. S. Virus inhibition of RIP3-dependent necrosis. *Cell Host Microbe* **7**, 302–313 (2010).
- Kelliher, M. A. *et al.* The death domain kinase RIP mediates the TNF-induced NF- κ B signal. *Immunity* **8**, 297–303 (1998).
- Cusson-Hermance, N., Khurana, S., Lee, T. H., Fitzgerald, K. A. & Kelliher, M. A. Rip1 mediates the Trif-dependent toll-like receptor 3- and 4-induced NF- κ B activation but does not contribute to interferon regulatory factor 3 activation. *J. Biol. Chem.* **280**, 36560–36566 (2005).
- Meylan, E. *et al.* RIP1 is an essential mediator of Toll-like receptor 3-induced NF- κ B activation. *Nature Immunol.* **5**, 503–507 (2004).
- Kawai, T. *et al.* IPS-1, an adaptor triggering RIG-I and Mda5-mediated type I interferon induction. *Nature Immunol.* **6**, 981–988 (2005).
- Rajput, A. *et al.* RIG-I RNA helicase activation of IRF3 transcription factor is negatively regulated by caspase-8-mediated cleavage of the RIP1 protein. *Immunity* **34**, 340–351 (2011).
- Christofferson, D. E., Li, Y. & Yuan, J. Control of life-or-death decisions by RIP1 kinase. *Annu. Rev. Physiol.* **76**, 129–150 (2014).
- Vanden Berghe, T., Linkermann, A., Jouan-Lanhout, S., Walczak, H. & Vandenabeele, P. Regulated necrosis: the expanding network of non-apoptotic cell death pathways. *Nature Rev. Mol. Cell Biol.* **15**, 135–147 (2014).
- Hill, D. A. & Artis, D. Intestinal bacteria and the regulation of immune cell homeostasis. *Annu. Rev. Immunol.* **28**, 623–667 (2010).
- Kaser, A., Zeissig, S. & Blumberg, R. S. Inflammatory bowel disease. *Annu. Rev. Immunol.* **28**, 573–621 (2010).
- Lee, T. H., Shank, J., Cusson, N. & Kelliher, M. A. The kinase activity of Rip1 is not required for tumor necrosis factor- α -induced I κ B kinase or p38 MAP kinase activation or for the ubiquitination of Rip1 by Traf2. *J. Biol. Chem.* **279**, 33185–33191 (2004).
- Greten, F. R. *et al.* IKK β links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. *Cell* **118**, 285–296 (2004).
- Nenci, A. *et al.* Epithelial NEMO links innate immunity to chronic intestinal inflammation. *Nature* **446**, 557–561 (2007).
- Steinbrecher, K. A., Harmel-Laws, E., Sitcheran, R. & Baldwin, A. S. Loss of epithelial RelA results in deregulated intestinal proliferative/apoptotic homeostasis and susceptibility to inflammation. *J. Immunol.* **180**, 2588–2599 (2008).
- Sasaki, Y. *et al.* Canonical NF- κ B activity, dispensable for B cell development, replaces BAFF-receptor signals and promotes B cell proliferation upon activation. *Immunity* **24**, 729–739 (2006).
- Vlantis, K. *et al.* Constitutive IKK2 activation in intestinal epithelial cells induces intestinal tumors in mice. *J. Clin. Invest.* **121**, 2781–2793 (2011).
- Gentile, I. E. *et al.* In TNF-stimulated cells, RIPK1 promotes cell survival by stabilizing TRAF2 and cIAP1, which limits induction of non-canonical NF- κ B and activation of caspase-8. *J. Biol. Chem.* **286**, 13282–13291 (2011).
- Kim, J. Y. *et al.* TNF α induced noncanonical NF- κ B activation is attenuated by RIP1 through stabilization of TRAF2. *J. Cell Sci.* **124**, 647–656 (2011).
- Piao, J. H. *et al.* Tumor necrosis factor receptor-associated factor (TRAF) 2 controls homeostasis of the colon to prevent spontaneous development of murine inflammatory bowel disease. *J. Biol. Chem.* **286**, 17879–17888 (2011).
- Varfolomeev, E. *et al.* IAP antagonists induce autubiquitination of c-IAPs, NF- κ B activation, and TNF α -dependent apoptosis. *Cell* **131**, 669–681 (2007).
- Vince, J. E. *et al.* IAP antagonists target cIAP1 to induce TNF α -dependent apoptosis. *Cell* **131**, 682–693 (2007).

Acknowledgements We are grateful to V. Dixit for *Ripk3*^{-/-}, D. Gumucio for Villin-Cre and S. Robine for Villin-CreER^{T2} mice. We thank C. Uthoff-Hachenberg, J. Buchholz, E. Mählberg, B. Kühnel, B. Hülser, P. Jankowski, S. Assenmacher and P. Scholl for technical assistance. M.P. acknowledges funding from the European Research Council (2012-ADG_20120314), the German Research Council (DFG; SFB670, SFB829, SPP1656), the European Commission (grants 223404 (Masterswitch) and 223151 (InflaCare)), the Deutsche Krebshilfe, the Else Kröner-Fresenius-Stiftung and the Helmholtz Alliance (PCCC). Research reported in this publication was also supported by the National Institute of Allergy and Infectious Diseases division of the National Institutes of Health under award RO1AI075118 to M.K.

Author Contributions M.D. together with K.V. performed and analysed the experiments related to the intestine and S.K. performed and analysed the experiments related to the skin. N.H. and M.K. designed and generated the targeting constructs for the *Ripk1*^{fl/fl} and *Ripk1*^{D138N/D138N} mice. A.P. performed the gene targeting in embryonic stem cells and generated the *Ripk1*^{fl/fl}, *Ripk1*^{D138N/D138N} and *Trif*^{fl/fl} mice. M.Z. contributed to the analysis of intestines from *Ripk1*^{-/-} neonates. C.K. and J.L. performed biochemical analysis of RIPK1-deficient MEFs, IECs and keratinocytes. C.E. performed FACS analysis of intestinal immune cells. T.C. performed qRT-PCR analysis. L.W. designed and tested the short guiding RNAs for CRISPR/Cas9-mediated targeting of *Mkl1*. P.K., M.B. and A.B. generated germ-free RIPK1^{IEC-KO} mice. M.P. coordinated the project and together with K.V., M.D. and S.K. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.P. (pasparakis@uni-koeln.de).

RIPK1 ensures intestinal homeostasis by protecting the epithelium against apoptosis

Nozomi Takahashi^{1,2}, Lars Vereecke^{1,2}, Mathieu J. M. Bertrand^{1,2}, Linde Duprez^{1,2}, Scott B. Berger³, Tatyana Divert^{1,2}, Amanda Gonçalves^{1,2,4}, Mozes Sze^{1,2}, Barbara Gilbert^{1,2}, Stephanie Kourula^{1,2}, Vera Goossens^{1,2}, Sylvie Lefebvre^{1,2}, Claudia Günther⁵, Christoph Becker⁵, John Bertin³, Peter J. Gough³, Wim Declercq^{1,2}, Geert van Loo^{1,2} & Peter Vandenabeele^{1,2,6}

Receptor interacting protein kinase 1 (RIPK1) has an essential role in the signalling triggered by death receptors and pattern recognition receptors^{1,2}. RIPK1 is believed to function as a node driving NF- κ B-mediated cell survival and inflammation as well as caspase-8 (CASP8)-dependent apoptotic or RIPK3/MLKL-dependent necroptotic cell death. The physiological relevance of this dual function has remained elusive because of the perinatal death of RIPK1 full knockout mice³. To circumvent this problem, we generated RIPK1 conditional knockout mice, and show that mice lacking RIPK1 in intestinal epithelial cells (IECs) spontaneously develop severe intestinal inflammation associated with IEC apoptosis leading to early death. This early lethality was rescued by antibiotic treatment, MYD88 deficiency or tumour-necrosis factor (TNF) receptor 1 deficiency, demonstrating the importance of commensal bacteria and TNF in the IEC *Ripk1* knockout phenotype. CASP8 deficiency, but not RIPK3 deficiency, rescued the inflammatory phenotype completely, indicating the indispensable role of RIPK1 in suppressing CASP8-dependent apoptosis but not RIPK3-dependent necroptosis in the intestine. RIPK1 kinase-dead knock-in mice did not exhibit any sign of inflammation, suggesting that RIPK1-mediated protection resides in its kinase-independent platform function. Depletion of RIPK1 in intestinal organoid cultures sensitized them to TNF-induced apoptosis, confirming the *in vivo* observations. Unexpectedly, TNF-mediated NF- κ B activation remained intact in these organoids. Our results demonstrate that RIPK1 is essential for survival of IECs, ensuring epithelial homeostasis by protecting the epithelium from CASP8-mediated IEC apoptosis independently of its kinase activity and NF- κ B activation.

RIPK1 is a serine/threonine kinase that is an upstream regulator of the cellular responses initiated by members of the death receptor families, such as TNF receptor (TNFR), and pattern recognition receptor super-families^{1,2}. Downstream of TNFR1 (also known as TNFRSF1A), RIPK1 regulates two opposed cellular fates: cell survival via NF- κ B-mediated gene induction, and cell death by induction of apoptosis or necroptosis. The kinase activity of RIPK1 is dispensable for its pro-survival function but essential for its killing potential^{4,5}. When CASP8 is absent or inhibited, RIPK1 triggers TNF-mediated necroptosis through a phosphorylation-driven cascade involving RIPK3 as a partner^{6–8} and MLKL as the downstream cell death executor^{9,10}. When inhibitor of apoptosis (IAP), transforming-growth-factor- β -activated kinase 1 or NF- κ B essential modulator are inhibited, RIPK1 mediates CASP8-dependent apoptosis^{11–14}. The mechanisms by which RIPK1 switches from a pro-survival to a pro-death molecule are poorly understood but probably depend on its ubiquitination state^{14–16}. The *in vivo* role of RIPK1 has been difficult to assess due to the perinatal mortality of full knockout mice³. Therefore, we generated a RIPK1 conditional knockout mouse line (*Ripk1*^{fl/fl}) by loxP technology (Extended Data Fig. 1a–c). *Ripk1*^{fl/fl} progeny of

Ripk1^{fl/fl} mice crossed with Cre-deleter mice died soon after birth, as reported³. *Ripk1* depletion was confirmed in primary mouse embryonic fibroblasts (MEFs) (Extended Data Fig. 1d). Upon TNF stimulation, *Ripk1*^{fl/fl} MEFs showed defective activation of mitogen-activated protein kinases (MAPKs) and canonical NF- κ B associated with apoptosis (Extended Data Fig. 2a, b).

Defects in the NF- κ B pathway resulting from deletion of NF- κ B essential modulator or transforming-growth-factor- β -activated kinase 1 in IECs cause massive apoptosis and gut inflammation resembling inflammatory bowel disease^{17,18}. However, RIPK3-dependent necroptosis was recently demonstrated to drive inflammatory-bowel-disease-like gut inflammation in mice with CASP8 or Fas-associated protein with death domain (FADD) deletion in IECs, identifying CASP8 and FADD as physiological suppressors of necroptosis^{19,20}. Since RIPK1 has a role in both NF- κ B activation and RIPK3-dependent cell death, we investigated the consequences of its deletion in IECs by crossing *Ripk1*^{fl/fl} mice with villin-Cre transgenic mice to give an IEC-specific *Ripk1* knockout (*Ripk1*^{IEC-KO}). The progeny were born at a normal Mendelian ratio, but their growth was severely retarded and about 70% of them died within 4 weeks (Fig. 1a, b and Extended Data Fig. 3a). DNA and protein analysis of various organs from the *Ripk1*^{IEC-KO} mice confirmed specific *Ripk1* deletion in the colon and small intestine (Extended Data Fig. 3b, c). The colon of *Ripk1*^{IEC-KO} mice displays progressive inflammation characterized by the presence of inflammatory infiltrates, aberrant crypts, and loss of goblet cells and tissue architecture (Fig. 1c and Extended Data Fig. 3d) accompanied by apoptosis, as shown by terminal deoxynucleotidyl transferase dUTP nick end labelling (TUNEL) and CASP3 activation (Fig. 1c and Extended Data Fig. 3e). When *Ripk1* was deleted in adult *Ripk1*^{fl/fl} villin-cre-ERT2 transgenic mice (an inducible IEC knockout, *Ripk1*^{IEC-KO}) by using 4-hydroxytamoxifen (4-OHT), massive intestinal cell death occurred and all the mice died. Thus, the phenotype of *Ripk1*^{IEC-KO} mice is not neonate specific (Fig. 1d, e). Surviving *Ripk1*^{IEC-KO} mice showed splenomegaly and colon shortening (Extended Data Fig. 4a), and colonoscopic analysis revealed lesions associated with inflammatory dysplasia (Fig. 1f), consistent with the histopathological findings (Extended Data Fig. 4b). The inflammation in the small intestine was also associated with a loss of tissue architecture and of mucus-producing goblet cells and antimicrobial peptide (AMP)-producing Paneth cells (Fig. 1g), resulting in physical barrier dysfunction and intestinal leakage, as shown by increased fluorescence in the serum of *Ripk1*^{IEC-KO} mice fed with FITC-labelled dextran (Fig. 1h). Accordingly, expression of the AMPs lysozyme P and cryptidin 1 was decreased, whereas TNF production was increased (Fig. 1i). As cell death, inflammation and mortality were associated with bacterial colonization shortly after birth, we hypothesized that deletion of RIPK1 sensitized IECs to TNF that is produced upon sensing of commensal bacteria. This would result in barrier disruption, commensal infiltration and severe gut inflammation,

¹VIB Inflammation Research Center, Technologiepark 927, B-9052 Ghent, Belgium. ²Department of Biomedical Molecular Biology, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium.

³Pattern Recognition Receptor Discovery Performance Unit, Immuno-inflammation Therapeutic Area, GlaxoSmithKline, Collegeville, Pennsylvania 19426, USA. ⁴VIB Bio Imaging Core Gent, Technologiepark 927, B-9052 Ghent, Belgium. ⁵Department of Medicine 1, Friedrich-Alexander-University, D-91054 Erlangen, Germany. ⁶Methusalem program, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium.

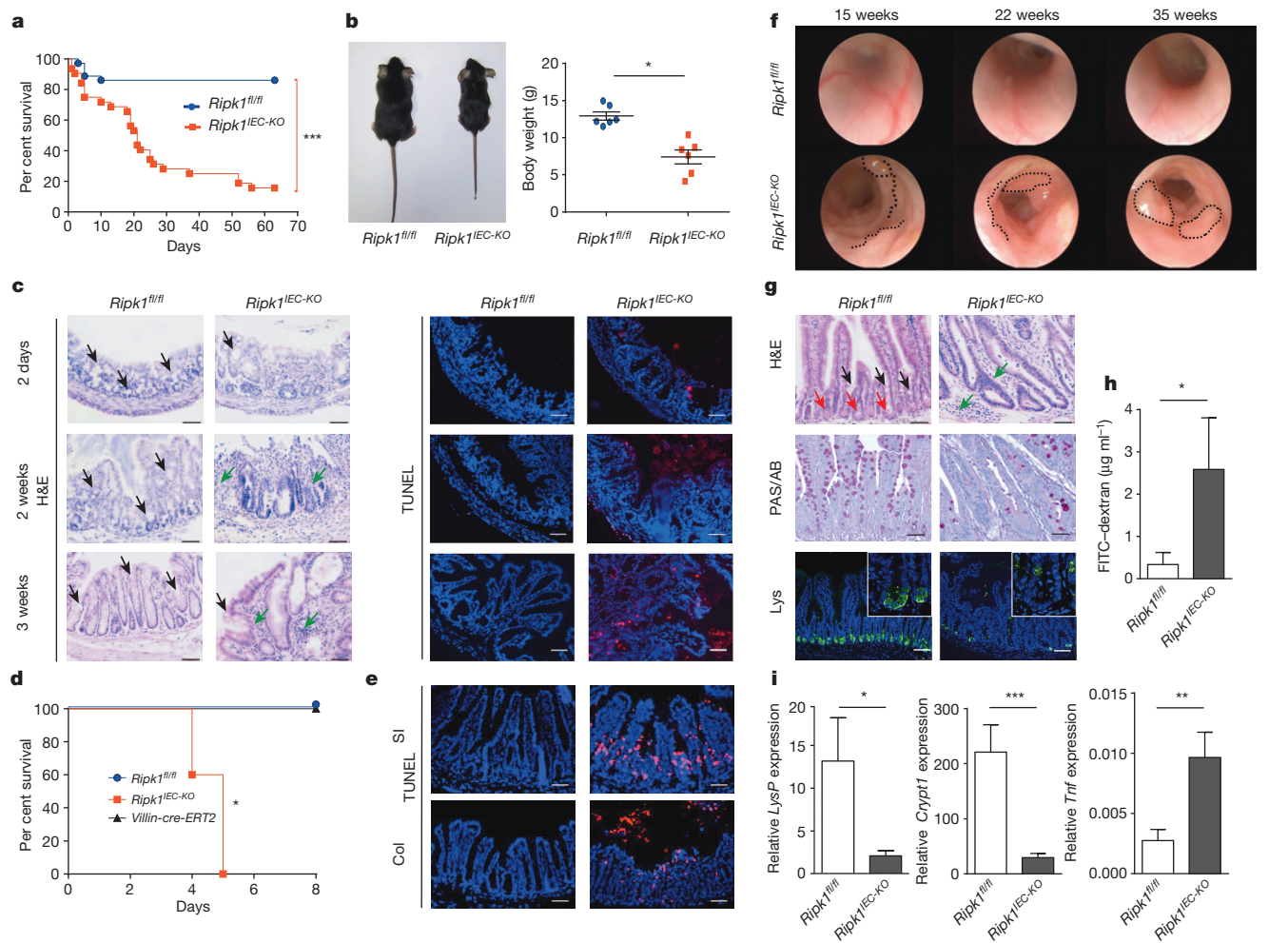


Figure 1 | Death and spontaneous gut inflammation in mice with RIPK1 deletion in IECs. **a**, Cumulative survival rate of *Ripk1^{IEC-KO}* mice ($n = 34$) compared to *Ripk1^{fl/fl}* mice ($n = 39$), *** $P < 0.0001$. **b**, Macroscopic features and body weight of *Ripk1^{IEC-KO}* (3 weeks, $n = 6$) and sex-matched littermate *Ripk1^{fl/fl}* mice ($n = 6$). * $P < 0.05$. Data represent mean \pm standard error of the mean (s.e.m.). **c**, Haematoxylin and eosin staining (H&E; left panels) and TUNEL assay (red; right panels) of colon sections from *Ripk1^{fl/fl}* and *Ripk1^{IEC-KO}* mice. Black arrows, goblet cells; green arrows, infiltrating leukocytes (scale bars, 50 μ m). **d**, Survival rates of *Ripk1^{fl/fl}* ($n = 3$), *Ripk1^{fl/fl} villin-cre-ERT2^{Tg}* (*Ripk1^{IEC-KO}*) ($n = 5$) and *villin-cre-ERT2^{Tg}* mice ($n = 2$) treated with 1 mg 4-OHT daily; * $P < 0.05$. **e**, TUNEL assay (red) in small intestine (SI) and colon (Col) of *Ripk1^{fl/fl}* and *Ripk1^{IEC-KO}* mice after three daily 4-OHT injections (scale bars, 25 μ m). Representative images of *Ripk1^{fl/fl}*

exacerbate TNF production and lead to the death of *Ripk1^{IEC-KO}* mice. We tested our hypothesis by using a cocktail of broad-spectrum antibiotics and by a genetic approach using mice deficient in MYD88—an essential adaptor protein in the signalling pathways activated downstream of toll-like receptors (TLRs) in response to pathogen-associated molecular pattern sensing—or mice deficient in TNFR1. Antibiotic treatment starting in the mother 2 weeks before birth of the offspring significantly improved the viability of *Ripk1^{IEC-KO}* mice and reduced inflammation and apoptosis in the majority of mice (Fig. 2a, b and Extended Data Fig. 5a, b). Body size, spleen size, colon length (Fig. 2c), number of Paneth cells (Extended Data Fig. 5b) and TNF level (Extended Data Fig. 5c) were normalized. The use of antibiotics also significantly prolonged the survival of adult *Ripk1^{IEC-KO}* mice following 4-OHT-induced *Ripk1* deletion (Extended Data Fig. 5d). In agreement, MYD88 significantly delayed the death of *Ripk1^{IEC-KO}* mice (Fig. 2d and Extended Data Fig. 6a). Nevertheless, histological signs of intestinal inflammation and apoptosis persisted in *Ripk1^{IEC-KO} Myd88^{-/-}* mice (Fig. 2e and

($n = 3$) and *Ripk1^{IEC-KO}* ($n = 3$) mice. **f**, Endoscopic images of *Ripk1^{IEC-KO}* and littermate *Ripk1^{fl/fl}* mice. Dotted lines show irregular surface and lesions. **g**, Haematoxylin and eosin and periodic acid–Schiff/Alcian blue (PAS/AB; goblet cells) staining and lysozyme (Lys) (Paneth cells) immunohistochemical staining in the small intestine. Black arrows, goblet cells; red arrows, Paneth cells; green arrows, infiltrating leukocytes (scale bars, 50 μ m). **h**, Intestinal permeability assay in *Ripk1^{IEC-KO}* ($n = 3$) and *Ripk1^{fl/fl}* mice ($n = 3$) using FITC-labelled dextran. Data represent mean \pm s.e.m. * $P < 0.05$. **i**, Quantitative real-time PCR (qPCR) analysis of antimicrobial peptides and TNF on RNA extracted from small intestine mucosa of *Ripk1^{IEC-KO}* ($n = 8$) and control *Ripk1^{fl/fl}* mice ($n = 8$). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Data represent mean \pm s.e.m. LysP, lysozyme P; Crypt1, cryptidin 1.

Extended Data Fig. 6c), indicating the involvement of additional MYD88-independent pathways. TNFR1 deficiency also significantly protected *Ripk1^{IEC-KO}* mice from death (Fig. 2f and Extended Data Fig. 6b), and although mild inflammation remained, apoptosis was inhibited in *Ripk1^{IEC-KO} Tnfr1^{-/-}* mice (Fig. 2g and Extended Data Fig. 6d). These observations indicate that bacterial colonization and subsequent production of TNF are crucial for amplification of apoptosis, contributing to intestinal inflammation and death.

We then examined whether blocking cell death in IECs would rescue the inflammatory phenotype by generating *Ripk1^{IEC-KO} Casp8^{IEC-KO}* and *Ripk1^{IEC-KO} Ripk3^{-/-}* double knockout mice. *Ripk1^{IEC-KO} Casp8^{IEC-KO}* mice were born at a normal Mendelian ratio, grew normally until adulthood and were macroscopically indistinguishable from *Ripk1^{fl/fl} Casp8^{fl/fl}* control mice (Fig. 3a and Extended Data Fig. 7a, b). In contrast *Ripk1^{IEC-KO} Ripk3^{-/-}* mice are phenotypically identical to *Ripk1^{IEC-KO}* mice (Fig. 3b and Extended Data Fig. 7c, d), excluding a role for RIPK3-dependent necroptosis⁶, inflammation²¹ or apoptosis¹³ in the *Ripk1^{IEC-KO}*

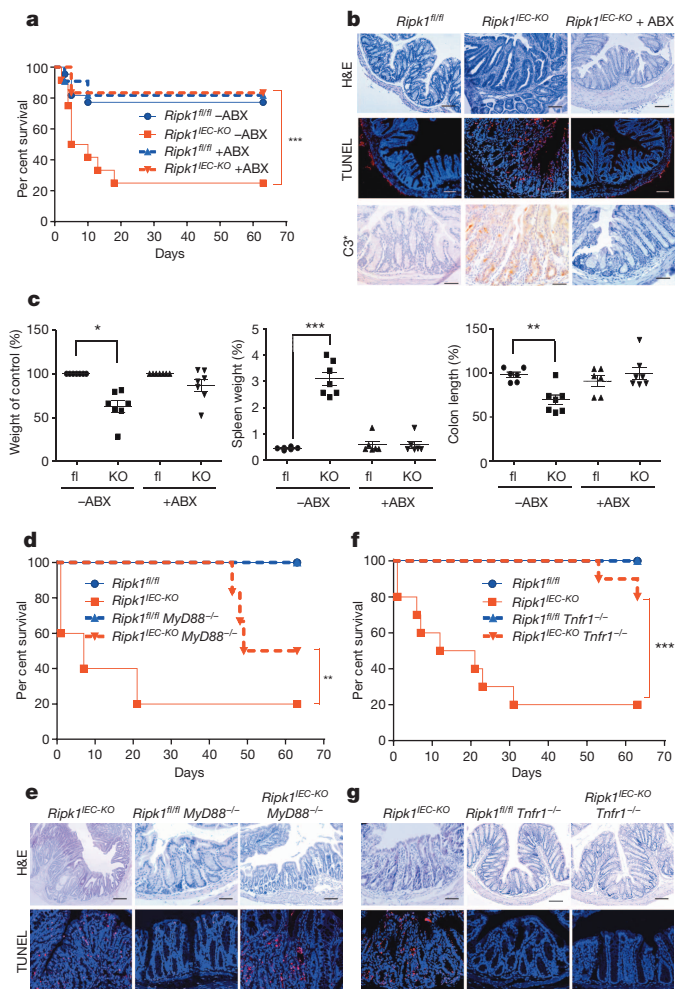


Figure 2 | *Rlpk1*^{IEC-KO} phenotype in conditions of antibiotic treatment, MYD88 deficiency, or TNFR1 deficiency. **a**, Breeding couples were treated with broad-spectrum antibiotics (+ABX) 2 weeks before the birth of offspring, and survival rates of *Rlpk1*^{fl/fl} (*n* = 18) and *Rlpk1*^{IEC-KO} mice (*n* = 12) were compared to untreated controls (–ABX). ****P* = 0.0008. **b**, Haematoxylin and eosin staining and TUNEL assay (red) and active CASP3 (C3*) immunohistochemistry (brown) of *Rlpk1*^{IEC-KO} and *Rlpk1*^{fl/fl} mice treated as in **a**. Representative image of *Rlpk1*^{fl/fl} (*n* = 3), untreated *Rlpk1*^{IEC-KO} (*n* = 7) and antibiotic-treated *Rlpk1*^{IEC-KO} mice (*n* = 7). **c**, Body size ratio (expressed as percent relative to littermate), spleen weight and colon length of 6–12-week-old *Rlpk1*^{IEC-KO} (*n* = 7) and littermate *Rlpk1*^{fl/fl} mice (*n* = 6) with or without antibiotics. **P* < 0.05, ***P* < 0.005, ****P* < 0.0001. Data represent mean ± s.e.m. **d**, Survival rate of *Rlpk1*^{IEC-KO} *Myd88*^{–/–} (*n* = 5) compared to *Rlpk1*^{IEC-KO} (*n* = 5), control *Rlpk1*^{fl/fl} (*n* = 5) and *Rlpk1*^{fl/fl} *Myd88*^{–/–} mice (*n* = 6). ****P* < 0.001. **e**, Haematoxylin and eosin staining (scale bars, 100 μm) and TUNEL assay (scale bars, 25 μm) of colon sections from *Rlpk1*^{fl/fl} *Myd88*^{–/–} and *Rlpk1*^{IEC-KO} *Myd88*^{–/–} mice. **f**, Survival rate of *Rlpk1*^{IEC-KO} *Tnfr1*^{–/–} (*n* = 9), *Rlpk1*^{IEC-KO} (*n* = 10), control *Rlpk1*^{fl/fl} (*n* = 11) and *Rlpk1*^{fl/fl} *Tnfr1*^{–/–} mice (*n* = 7). ****P* < 0.0001. **g**, Haematoxylin and eosin staining (scale bars, 100 μm) and TUNEL assay (scale bars, 25 μm) of colon sections from *Rlpk1*^{fl/fl} *Tnfr1*^{–/–} and *Rlpk1*^{IEC-KO} *Tnfr1*^{–/–} mice.

phenotype. Histopathological analysis of the colons revealed complete protection from apoptosis and inflammation by concomitant deletion of *Casp8* but not of *Ripk3* (Fig. 3c). This demonstrates that *Ripk1*^{IEC-KO} mice develop lethal inflammation due to excessive CASP8-dependent apoptosis of IECs, identifying RIPK1 as an apoptosis suppressor in the intestine. Complete rescue by deletion of CASP8 but not by TNFR1 deletion suggests that stimuli other than TNF might account for IEC apoptosis and inflammation and that other death-inducing stimuli converge at CASP8. However, the spontaneous ileitis caused by uncontrolled necroptosis in *Casp8*^{IEC-KO} mice, as reported earlier²⁰, persisted in *Ripk1*^{IEC-KO}

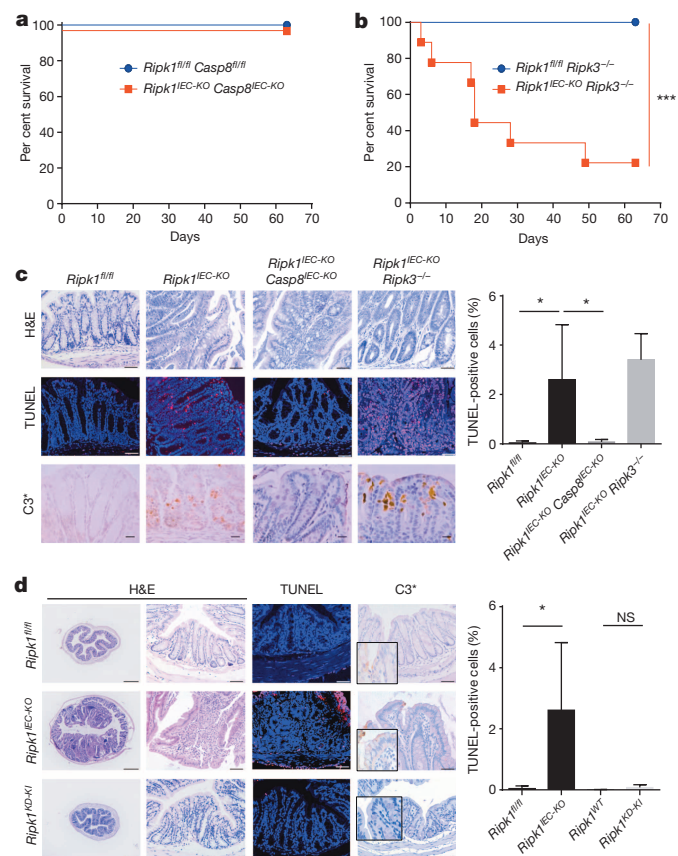


Figure 3 | Lethality of *Ripk1*^{IEC-KO} in mice requires *Casp8* but not RIPK3 or RIPK1 kinase activity. **a**, Survival rate of *Ripk1*^{IEC-KO} *Casp8*^{IEC-KO} mice (*n* = 6) and control *Ripk1*^{fl/fl} *Casp8*^{fl/fl} mice (*n* = 8). **b**, Survival rate of *Ripk1*^{IEC-KO} *Ripk3*^{–/–} mice (*n* = 9) and control *Ripk1*^{fl/fl} *Ripk3*^{–/–} mice (*n* = 9). ****P* < 0.0001. **c**, Histopathological analysis of colons of *Ripk1*^{fl/fl} (*n* = 4), *Ripk1*^{IEC-KO} (*n* = 4), *Ripk1*^{IEC-KO} *Casp8*^{IEC-KO} (*n* = 4) and *Ripk1*^{IEC-KO} *Ripk3*^{–/–} mice (*n* = 3) by haematoxylin and eosin staining and TUNEL assay (red; quantification in the right panel), and active CASP3 immunohistochemistry (C3*; brown). Scale bars, 50 μm. **d**, Histopathological analysis of colon samples from 12-week-old *Ripk1*^{fl/fl} (*n* = 3), *Ripk1*^{IEC-KO} (*n* = 3) and RIPK1-kinase-dead knock-in mice (*Ripk1*^{KD-KI}, *n* = 5) by haematoxylin and eosin staining and TUNEL assay (red; quantification in the right panel), and active CASP3 immunohistochemistry (C3*; brown). Scale bars, 50 μm. **P* < 0.05. Data represent mean ± s.e.m. WT, wild type; NS, not significant.

Casp8^{IEC-KO} mice, indicating that absence of RIPK1 does not protect *Casp8*^{IEC-KO} mice against necroptosis in the ileum (Extended Data Fig. 7e, f).

Because RIPK1 mediates both kinase-dependent and -independent functions^{1,2}, we evaluated the requirement for its enzymatic activity by characterizing a RIPK1 kinase-dead knock-in mouse line (*Ripk1*^{KD-KI})^{4,5}. Unlike the *Ripk1*^{–/–} and *Ripk1*^{IEC-KO} mice, *Ripk1*^{KD-KI} mice were viable and grew normally into adults^{4,5}. The colon (Fig. 3d) and small intestine (Extended Data Fig. 8) of *Ripk1*^{KD-KI} mice showed no sign of inflammation or apoptosis. This indicates that the *Ripk1*^{IEC-KO} phenotype is caused by loss of the RIPK1 platform-mediated function, as the kinase activity of RIPK1 is totally dispensable for the maintenance of intestinal homeostasis.

To investigate the mechanism of apoptosis in IECs further, we established intestinal organoid cultures²². Because crypt cells from *Ripk1*^{IEC-KO} mice failed to grow into organoids (data not shown), we depleted RIPK1 in organoids derived from *Ripk1*^{IEC-KO} mice by *in vitro* 4-OHT treatment. In accordance with our observations that death of IECs in *Ripk1*^{IEC-KO} mice is associated with increased *Tnf* mRNA levels in intestinal tissue (Fig. 1i), and that absence of TNFR1 rescues perinatal

death (Fig. 2f, g), we found that TNF treatment induced massive apoptosis in the *Ripk1*-depleted organoids but not in the control ones (Fig. 4a and Supplementary Videos 1–4), demonstrating enhanced sensitivity towards TNF as compared to control. In agreement, CASP8 and CASP3 were activated only in lysates from *Ripk1*-depleted cultures (Fig. 4b). Reportedly, RIPK1 mediates TNF-dependent canonical NF- κ B activation independently of its kinase activity, and the sensitivity of the RIPK1-deficient MEFs to TNF-mediated apoptosis has been explained, at least in part, by defective induction of NF- κ B-dependent pro-survival genes³ (Extended Data Fig. 2). Notably, the sensitization of *Ripk1*-depleted organoids to TNF-induced apoptosis was not associated with defects in canonical NF- κ B activation. I κ B α phosphorylation and degradation occurred with similar efficiency and kinetics in both cultures regardless of *Ripk1* depletion (Fig. 4b). Intact NF- κ B activation was further confirmed by analysing the induction of NF- κ B-dependent genes in the organoid culture. Expression of these target genes did not decrease, while *Ripk1* depletion by 4-OHT treatment exceeded 90% (Fig. 4c). Finally, we confirm *in vitro* that the sensitization of RIPK1-deficient IECs to apoptosis is independent of the kinase function (Fig. 3d), as organoids derived from *Ripk1*^{KD-KI} mice were not sensitized to TNF (Fig. 4d–f and Supplementary Video 5). RIPK1-depleted organoids eventually died after 4 days without exogenous stimuli, confirming an essential role of RIPK1 for survival of IECs (Extended Data Fig. 9a and Supplementary Videos 6 and 7). Whether this TNF-independent cell death or additional death receptor ligands contribute to *in vivo* phenotype remains unknown, but administration of prototype pathogen-associated molecular patterns such as lipopolysaccharide and lipoteichoic acid do not directly induce cell death of intestinal organoids (Extended Data Fig. 9b, c).

In conclusion, we demonstrate that RIPK1 has a crucial kinase-independent role in protecting IECs from CASP8-dependent apoptosis by maintaining intestinal homeostasis. Low constitutive levels of TNF secreted in the mucosa may be sufficient to kill IECs lacking RIPK1 and partially disrupt the intestinal barrier, allowing commensal bacteria to infiltrate and activate innate immunity. The consequent increase in TNF fuels a cycle of IEC apoptosis and inflammation in mice lacking homeostatic control by RIPK1. Notably, RIPK1 deficiency sensitizes IECs to TNF-mediated apoptosis independently of any defect in NF- κ B activation, whereas sensitization of RIPK1-deficient MEFs was associated with defective NF- κ B, revealing an unexpected pro-survival role of RIPK1 downstream of TNF in IECs. Of note, one study suggested that RIPK1 protects MEFs from TNF-mediated CASP8-dependent apoptosis by promoting cFLIP stability independently of canonical NF- κ B activation²³ (Extended Data Fig. 10a). We observed a modest decrease of cFLIP level in RIPK1-depleted IECs stimulated with TNF, but the biological relevance of this observation remains unclear (Extended Data Fig. 10b). This platform-dependent anti-apoptotic RIPK1 function might exist in several cell types, as RIPK1 depletion in L929 cells also sensitized them to TNF-induced CASP8-dependent apoptosis²⁴. Our results support a novel concept that CASP8 keeps RIPK1 and RIPK3 necroptotic signalling in check^{19,25,26}, while RIPK1 protects cells from excessive CASP8 activation in response to mild triggering with death receptor ligands or cellular stress. This means that the RIPK1–CASP8 node is the crucial decision point in cell death or survival and explains why cells can switch so easily from apoptosis to necroptosis and vice versa, when either cell death pathway is blocked^{24,27}. In contrast to three recent papers reporting that rescuing the lethality of RIPK1 full knockout requires simultaneous block of CASP8-dependent apoptosis and RIPK3-dependent necroptosis^{28–30}, we found that blocking apoptosis alone rescues intestinal inflammation. This finding indicates that the default cell death mode is tissue specific. Together, our results emphasize that RIPK1 is an essential survival factor in intestinal homeostasis. An accompanying paper³¹ presents similar data. Notably, different approaches to rescue the phenotype of *Ripk1*^{IEC-KO} mice by either ablating CASP8 (our study) or FADD (ref. 31) prevented IEC apoptosis caused by RIPK1 deficiency. However, necroptosis in IECs caused by deficiency of CASP8 or FADD cannot be prevented by loss of RIPK1, suggesting RIPK1-independent

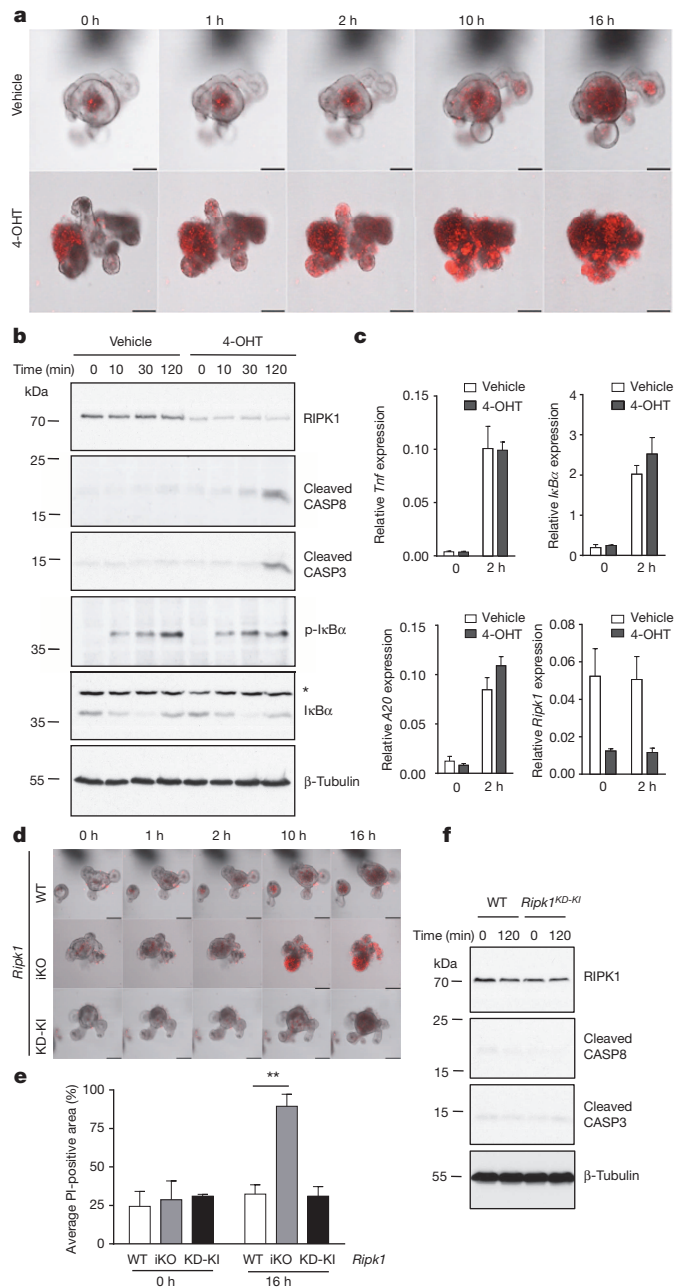


Figure 4 | *Ripk1*^{IEC-KO} intestinal organoids are sensitized to TNF-induced apoptosis independently of NF- κ B and kinase activity. Intestinal organoids were derived from *Ripk1*^{IEC-KO} mice; RIPK1 deletion was induced *in vitro* by 4-OHT treatment (200 nM) for 20–24 h after which organoids were treated with recombinant mouse TNF (mTNF). **a**, Time-lapse confocal microscopic imaging of propidium iodide (PI)-stained organoids. Scale bars, 100 μ m. **b**, Western blot analysis of organoid cultures at different time points after mTNF administration. p-I κ B α , phosphorylated I κ B α . **c**, qPCR analysis of organoids stimulated with mTNF for 2 h ($n = 3$ biological replicates). Data represent mean \pm s.e.m. A20, TNF α -induced protein 3. **d**, Intestinal organoids derived from *Ripk1*^{fl/fl} (wild type, WT), *Ripk1*^{IEC-KO} (iKO) and *Ripk1*^{KD-KI} mice were treated with mTNF as in **a**. **e**, Image analysis of the per cent of propidium iodide-positive nuclei as in **d** ($n = 3$, biological triplicates). ****** $P < 0.01$. Data represent mean \pm s.e.m. WT, wild type; iKO, inducible knockout. **f**, Western blot analysis of organoid cultures after mTNF administration as in **d**.

necroptosis. Disease kinetics differ between the two studies and the pathology of our *Ripk1*^{IEC-KO} mice is milder and consequently may be better protected by rescue experiments using antibiotics or genetic deletions such as *Tnfr1*^{-/-} or *Myd88*^{-/-}. The reasons for these differences in severity of disease symptoms are not clear, but might be explained by

small variances in the genetic makeup of the mice strains used in different laboratory settings.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 February; accepted 23 July 2014.

- Declercq, W., Vanden Berghe, T. & Vandenabeele, P. RIP kinases at the crossroads of cell death and survival. *Cell* **138**, 229–232 (2009).
- Ofengeim, D. & Yuan, J. Regulation of RIP1 kinase signalling at the crossroads of inflammation and cell death. *Nature Rev. Mol. Cell Biol.* **14**, 727–736 (2013).
- Kelliher, M. A. *et al.* The death domain kinase RIP mediates the TNF-induced NF- κ B signal. *Immunity* **8**, 297–303 (1998).
- Berger, S. B. *et al.* Cutting edge: RIP1 kinase activity is dispensable for normal development but is a key regulator of inflammation in SHARPIN-deficient mice. *J. Immunol.* **192**, 5476–5480 (2014).
- Newton, K. *et al.* Activity of protein kinase RIPK3 determines whether cells die by necroptosis or apoptosis. *Science* **343**, 1357–1360 (2014).
- Cho, Y. S. *et al.* Phosphorylation-driven assembly of the RIP1–RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell* **137**, 1112–1123 (2009).
- He, S. *et al.* Receptor interacting protein kinase-3 determines cellular necrotic response to TNF- α . *Cell* **137**, 1100–1111 (2009).
- Zhang, D. W. *et al.* RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* **325**, 332–336 (2009).
- Murphy, J. M. *et al.* The pseudokinase MLKL mediates necroptosis via a molecular switch mechanism. *Immunity* **39**, 443–453 (2013).
- Sun, L. *et al.* Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell* **148**, 213–227 (2012).
- Tenev, T. *et al.* The ripoptosome, a signaling platform that assembles in response to genotoxic stress and loss of IAPs. *Mol. Cell* **43**, 432–448 (2011).
- Feoktistova, M. *et al.* cIAPs block ripoptosome formation, a RIP1/caspase-8 containing intracellular cell death complex differentially regulated by cFLIP isoforms. *Mol. Cell* **43**, 449–463 (2011).
- Dondelinger, Y. *et al.* RIPK3 contributes to TNFR1-mediated RIPK1 kinase-dependent apoptosis in conditions of cIAP1/2 depletion or TAK1 kinase inhibition. *Cell Death Differ.* **20**, 1381–1392 (2013).
- Bertrand, M. J. *et al.* cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Mol. Cell* **30**, 689–700 (2008).
- Ea, C. K., Deng, L., Xia, Z. P., Pineda, G. & Chen, Z. J. Activation of IKK by TNF α requires site-specific ubiquitination of RIP1 and polyubiquitin binding by NEMO. *Mol. Cell* **22**, 245–257 (2006).
- O'Donnell, M. A., Legarda-Addison, D., Skountzos, P., Yeh, W. C. & Ting, A. T. Ubiquitination of RIP1 regulates an NF- κ B-independent cell-death switch in TNF signaling. *Curr. Biol.* **17**, 418–424 (2007).
- Kajino-Sakamoto, R. *et al.* TGF- β -activated kinase 1 signaling maintains intestinal integrity by preventing accumulation of reactive oxygen species in the intestinal epithelium. *J. Immunol.* **185**, 4729–4737 (2010).
- Nenci, A. *et al.* Epithelial NEMO links innate immunity to chronic intestinal inflammation. *Nature* **446**, 557–561 (2007).
- Günther, C. *et al.* Caspase-8 regulates TNF- α -induced epithelial necroptosis and terminal ileitis. *Nature* **477**, 335–339 (2011).
- Welz, P. S. *et al.* FADD prevents RIP3-mediated epithelial cell necrosis and chronic intestinal inflammation. *Nature* **477**, 330–334 (2011).
- Vince, J. E. *et al.* Inhibitor of apoptosis proteins limit RIP3 kinase-dependent interleukin-1 activation. *Immunity* **36**, 215–227 (2012).
- Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
- Gentile, I. E. *et al.* In TNF-stimulated cells, RIPK1 promotes cell survival by stabilizing TRAF2 and cIAP1, which limits induction of non-canonical NF- κ B and activation of caspase-8. *J. Biol. Chem.* **286**, 13282–13291 (2011).
- Vanlangenakker, N. *et al.* cIAP1 and TAK1 protect cells from TNF-induced necrosis by preventing RIP1/RIP3-dependent reactive oxygen species production. *Cell Death Differ.* **18**, 656–665 (2011).
- Oberst, A. *et al.* Catalytic activity of the caspase-8–FLIP_L complex inhibits RIPK3-dependent necrosis. *Nature* **471**, 363–367 (2011).
- Kaiser, W. J. *et al.* RIP3 mediates the embryonic lethality of caspase-8-deficient mice. *Nature* **471**, 368–372 (2011).
- Moujalled, D. M. *et al.* In mouse embryonic fibroblasts, neither caspase-8 nor cellular FLICE-inhibitory protein (FLIP) is necessary for TNF to activate NF- κ B, but caspase-8 is required for TNF to cause cell death, and induction of FLIP by NF- κ B is required to prevent it. *Cell Death Differ.* **19**, 808–815 (2012).
- Dillon, C. P. *et al.* RIPK1 blocks early postnatal lethality mediated by caspase-8 and RIPK3. *Cell* **157**, 1189–1202 (2014).
- Rickard, J. A. *et al.* RIPK1 regulates RIPK3–MLKL-driven systemic inflammation and emergency hematopoiesis. *Cell* **157**, 1175–1188 (2014).
- Kaiser, W. J. *et al.* RIP1 suppresses innate immune necrotic as well as apoptotic cell death during mammalian parturition. *Proc. Natl Acad. Sci. USA* **111**, 7753–7758 (2014).
- Dannopp, M. *et al.* RIPK1 maintains epithelial homeostasis by inhibiting apoptosis and necroptosis. *Nature* <http://dx.doi.org/10.1038/nature13608> (this issue).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the Transgenic Mice Core Facility (IRC, VIB-UGent, Ghent) for their assistance. Villin-Cre, villin-Cre-ERT2 and RIPK3 knockout mice were provided by D. Gumucio (University of Michigan), S. Robine (Paris, France) and K. Newton and V. Dixit (Genentech), respectively. We thank A. Bredan for editing the manuscript, and P. De Bleser and M. Vuylsteke for statistical analysis. P.V. is senior full professor at Ghent University and holder of a Methusalem grant. L.V. is holder of a fellowship from Research Foundation Flanders (FWO). W.D. and G.V.L. have research professor positions at Ghent University. M.B. has a tenure track position in the Multidisciplinary Research Program of Ghent University (GROUP-ID). N.T. and V.G. are paid by the Methusalem grant, S.L. by a VIB grant and S.K. by a grant from the FWO. G.V.L. was supported by an FWO Odysseus Grant and by research grants from FWO, Foundation against Cancer and the Queen Elisabeth Medical foundation. Research in P.V.'s group is supported by Belgian grants (Interuniversity Attraction Poles, IAP 7/32), Flemish grants (Research Foundation Flanders, FWO G.0875.11, FWO G.0973.11, FWO G.0A45.12N, FWO G.0787.13N, G.0544.11N, GOC3114N and Methusalem grant BOF09/01M00709), Ghent University grants (MRP, GROUP-ID consortium), grant from the Foundation against Cancer (F94 and 2010-162) and grants from VIB. C.B. and C.G. received funding from the IZKF of the FAU Erlangen-Nürnberg and the DFG within the projects SPP1656, BE3686/2 and SFB796.

Author Contributions N.T., L.V., M.J.M.B., W.D., G.v.L. and P.V. designed the study. N.T., L.V., M.J.M.B., L.D., T.D., M.S., S.K., V.G., S.L., C.G. and B.G. performed the experiments. N.T., L.V., M.J.M.B., A.G., W.D., G.v.L. and P.V. analysed the data. N.T., M.J.M.B., W.D., G.v.L. and P.V. wrote the manuscript. S.B.B., C.B., J.B. and P.J.G. provided reagents and scientific insight.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.V. (Peter.Vandenabeele@irc.vib-ugent.be).

Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia

Serkan Kir¹, James P. White¹, Sandra Kleiner¹, Lawrence Kazak¹, Paul Cohen¹, Vickie E. Baracos² & Bruce M. Spiegelman¹

Cachexia is a wasting disorder of adipose and skeletal muscle tissues that leads to profound weight loss and frailty. About half of all cancer patients suffer from cachexia, which impairs quality of life, limits cancer therapy and decreases survival. One key characteristic of cachexia is higher resting energy expenditure levels than in healthy individuals, which has been linked to greater thermogenesis by brown fat^{1–6}. How tumours induce brown fat activity is unknown. Here, using a Lewis lung carcinoma model of cancer cachexia, we show that tumour-derived parathyroid-hormone-related protein (PTHrP) has an important role in wasting, through driving the expression of genes involved in thermogenesis in adipose tissues. Neutralization of PTHrP in tumour-bearing mice blocked adipose tissue browning and the loss of muscle mass and strength. Our results demonstrate that PTHrP mediates energy wasting in fat tissues and contributes to the broader aspects of cancer cachexia. Thus, neutralization of PTHrP might hold promise for ameliorating cancer cachexia and improving patient survival.

Cachexia is a wasting disorder that accompanies many chronic diseases such as cancer and congestive heart failure. The hallmarks of cachexia are weight loss with atrophy of fat and skeletal muscle⁷. Importantly, cachexia is more than just anorexia^{7–9}. Cachectic patients may ingest less food, but they are also in a state of negative energy balance that cannot be corrected by nutritional supplementation^{8,9}. As a negative risk factor for cancer survival, cachexia causes frailty in patients and often prevents them from undergoing further therapies. While cachexia can lessen with the shrinkage of tumours, there are currently few effective therapies for cancer cachexia^{10,11}.

The molecular basis of cachexia is poorly understood. Elevated cytokine levels have been observed in cachectic patients, but anti-cytokine therapies are ineffective^{10,11}. The activation of brown fat has been described in rodent models of cachexia and in certain cachectic patients^{1–6}. Brown fat dissipates chemical energy in the form of heat and, hence, could be involved in the negative energy balance. How tumours induce thermogenesis in brown fat cells and how this might relate to the wasting of fat and skeletal muscle is unknown. In addition, white fat depots contain pockets of uncoupling protein 1 (UCP1)-expressing multilocular cells, called beige (or brite) cells, that can be stimulated on exposure to cold or other stimuli via a process termed browning. Recent studies have characterized beige cells at a molecular level¹² and have demonstrated a requirement for the transcriptional coregulator PRDM16 in the browning process¹³. To understand the molecular basis of how tumours stimulate brown or beige cells, we used a murine model of lung cancer that is accompanied by cachexia: Lewis lung carcinoma (LLC).

LLC cells readily form tumours and lead to cachexia in syngeneic C57BL/6 mice. LLC-tumour-bearing mice demonstrated hypermetabolism: their rate of oxygen consumption was significantly higher than that of non-tumour-bearing mice (Fig. 1a). However, this increase was not due to increased physical activity or reduced food intake (Extended Data Fig. 1a, b). The lower carbon dioxide production relative to oxygen consumption (that is, the lowered respiratory exchange ratio) of LLC-tumour-bearing mice indicated that fat was the preferred source of fuel (Extended Data Fig. 1c, d). Additionally, more heat production was observed in the cachectic mice (Extended Data Fig. 1e). All tumour-bearing

mice lost weight over 3 weeks (Fig. 1b), including wasting of adipose tissues and skeletal muscle (Fig. 1c). Inguinal white adipose tissue (iWAT), a form of subcutaneous fat, wasted faster than epididymal fat (eWAT), a visceral white fat depot. In response to tumour growth, both white fat depots (eWAT and iWAT) and interscapular brown adipose tissue (iBAT) exhibited higher expression of genes involved in thermogenesis: *Ucp1*, *Dio2* and *Pgc1a* (Fig. 1d–f). An increase in the expression of certain other genes involved in energy metabolism, including glucose transporters and β -oxidation and lipolytic enzymes¹⁴, was also observed (Fig. 1d–f). LLC tumours also induced the skeletal muscle expression of atrophy-associated genes (Fig. 1g), including the genes encoding myostatin (*Mstn*), atrogin 1 (also known as *Fbxo32*) and MuRF1 (also known as *Trim63*).

To examine the contribution of adipose tissue browning to cachexia, we studied fat-specific *Prdm16*-deficient mice, which have a dramatically reduced thermogenic potential and are resistant to browning¹³. LLC-tumour-induced adipose tissue wasting was significantly inhibited in *Prdm16*-knockout mice, without a change in tumour size or the loss of muscle mass (Extended Data Fig. 2a, b). The cachexia-associated induction of *Ucp1* and *Dio2* expression was impaired in the knockout mice, while the increase in expression of the other genes involved in energy metabolism was unaffected (Extended Data Fig. 2c–f). These results indicate a crucial role for the beige fat thermogenesis pathway in adipose tissue wasting.

We examined whether LLC-tumour-derived factors may induce thermogenic gene expression in fat tissues. The treatment of primary white adipose cells with LLC-cell-conditioned medium stimulated the expression of *Ucp1* and *Dio2* (Fig. 2a). We then filtered the LLC-cell-conditioned medium through a 3-kDa cut-off membrane and tested the response of primary adipocytes to each fraction (Fig. 2b). All of the thermogenic activity was retained in the filter-concentrated fraction, leaving none in the flow-through. This finding suggests that the thermogenesis-inducing activity most probably originates from a macromolecule.

To identify thermogenic factors, we used a combination of cell cloning and gene expression profiling. We first established single-cell clones from the heterogeneous LLC cells. Thirty subclones were generated and used to produce conditioned media, the activity of which was then tested on primary adipocytes. We observed a range of *Ucp1*-inducing activity by conditioned media from these subclones (Fig. 2c), and we chose eight clones that demonstrated either more or less thermogenic activity than the parent cells (Extended Data Fig. 3a). We analysed the global gene expression profiles of the 8 subclones using microarrays and generated a list of secreted proteins for which the corresponding messenger RNA levels were higher in the more thermogenic subclones (Fig. 2d). We tested the effects of these secreted proteins on primary adipocytes and found that several members of the epidermal growth factor (EGF) family (Extended Data Fig. 3b), including betacellulin (BTC), heparin-binding EGF-like growth factor (HBEGF) and epiregulin (EREG), as well as PTHrP (encoded by *Pthlh*), induced *Ucp1* expression (Fig. 2e).

To determine the contribution of BTC, EREG and HBEGF to the *Ucp1*-inducing activity found in LLC-cell-conditioned medium, we used an

¹Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02215, USA. ²Department of Oncology, Division of Palliative Care Medicine, University of Alberta, Edmonton T6G 1Z2, Canada.

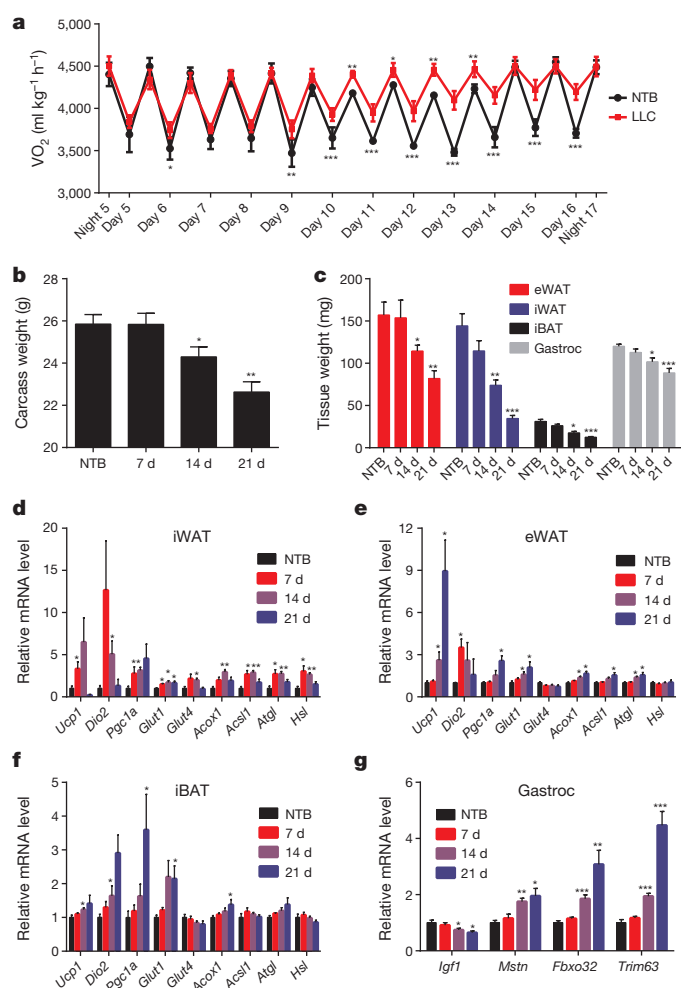


Figure 1 | LLC tumours cause adipose tissue browning and cachexia. Mice were inoculated with LLC cells and monitored for up to 21 days ($n = 6$ for each group). Four days after inoculation, a cohort of mice ($n = 6$ for the non-tumour-bearing (NTB) group and $n = 9$ for the LLC group) was placed into metabolic cages to measure oxygen consumption (VO_2) (a). Carcass weight (calculated by subtracting tumour weight from the total weight) (b) and the weight of fat and muscle tissues (c) were measured. The mRNA levels in iWAT (d), eWAT (e), iBAT (f) and the gastrocnemius muscle (Gastroc) (g) were determined by quantitative reverse transcription PCR (qRT-PCR). The values are mean \pm s.e.m. Statistical analysis was conducted using the two-tailed t -test. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$ compared with the NTB group. d, day.

inhibitor of EGF receptor (EGFR) signalling. Since the three EGF family proteins signal through the receptors EGFR and ERBB4, we tested the effects of AST-1306, a specific protein kinase inhibitor that targets both receptors, on primary adipocytes. AST-1306 completely blocked *Ucp1* mRNA induction when fat cells were treated in combination with BTC, EREG or HBEGF (Extended Data Fig. 3c). By contrast, the browning activity present in LLC-cell-conditioned medium was largely unaffected by AST-1306 treatment, suggesting that these EGF family proteins do not have a major quantitative role in this process (Fig. 3a).

PTHrP is processed into at least three peptide products. We treated primary white and brown adipocytes with these peptides and PTH(1–34) (a PTH-derived peptide containing amino acid residues 1 to 34), which has sequence and structural similarity to PTHrP(1–34). Both PTH(1–34) and PTHrP(1–34), which signal through the same receptor (the PTH and PTHrP receptor (PTHR)), stimulated thermogenic gene expression (Fig. 3b and Extended Data Fig. 3d). For simplicity, hereafter, the 1–34 peptides will be referred to as PTHrP and PTH. Time-course and dose-response analysis of PTHrP treatment revealed that this peptide

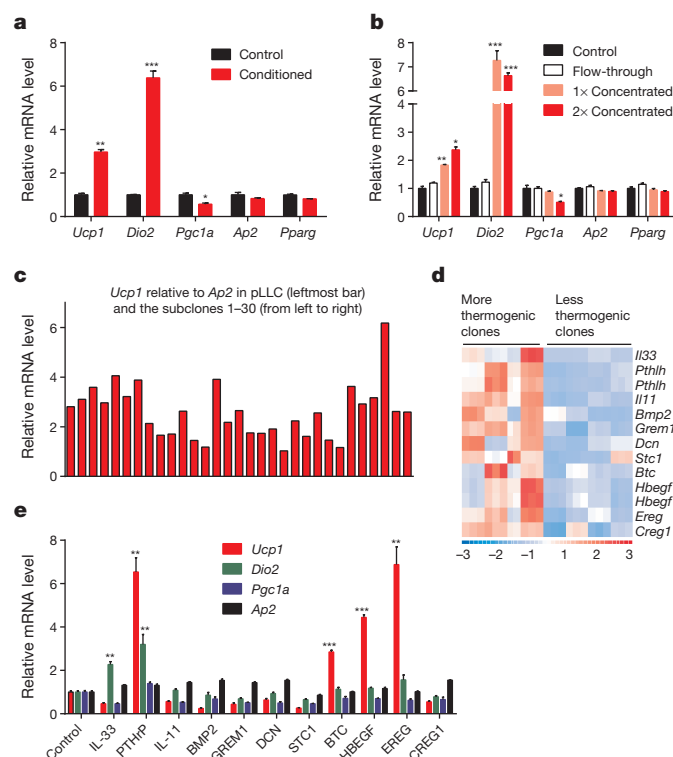


Figure 2 | LLC-cell-conditioned medium stimulates thermogenic gene expression in fat cells. a–c, Parent LLC cells or subclones were cultured in serum-free medium for 24 h. Primary adipocytes ($n = 3$ in a and b) were treated for 24 h with LLC-cell-conditioned medium (a and c) or filtered fractions (b). d, Gene expression heat map of the genes encoding the secreted factors identified in the microarray analysis. e, Primary adipocytes were treated with the indicated proteins ($1 \mu\text{g ml}^{-1}$) for 24 h ($n = 3$). The mRNA levels corresponding to the indicated genes were determined by qRT-PCR. *Ap2* (also known as *Fabp4*) and *Pparg* are controls for adipocyte differentiation. The values are mean \pm s.e.m. Statistical analysis was conducted using the two-tailed t -test. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$, compared with the control group.

potently induces *Ucp1* and *Dio2* mRNA (by up to 200- and 20-fold, respectively), comparable to noradrenaline, the classic thermogenic catecholamine produced by the sympathetic nervous system (Fig. 3c and Extended Data Fig. 4). We next investigated how PTHrP affects UCP1 protein levels and uncoupled respiration. Primary white and brown adipocytes were treated with noradrenaline, PTHrP or PTH. Remarkably, all three factors comparably increased UCP1 protein levels and significantly raised oxygen consumption, including basal, uncoupled and maximal respiration (Extended Data Fig. 5a–c). These results show that the PTH and PTHrP pathway can be a robust regulator of cellular respiration.

PTHR is a G-protein-coupled receptor that activates the cyclic-AMP-dependent protein kinase (PKA) pathway¹⁵. Noradrenaline also signals through β -adrenergic receptors that are G-protein-coupled receptors and induce PKA signalling, thereby promoting thermogenic gene expression. The treatment of primary white and brown adipocytes with noradrenaline, PTHrP or PTH stimulated the phosphorylation of the PKA substrates CREB and HSL. A selective PKA inhibitor, H89, blocked PKA signalling, which also completely inhibited transcriptional regulation by all three factors (noradrenaline, PTHrP and PTH) (Extended Data Fig. 5d–g). These results strongly suggest that both PTHrP and PTH require the PKA pathway to mediate their effects on *Ucp1* and *Dio2* transcription and therefore have a signalling mechanism in common with the β -adrenergic pathway.

Next, a neutralizing antibody specific for PTHrP was used to investigate whether PTHrP is involved in LLC-cell-induced thermogenic gene expression. This antibody was capable of completely inhibiting

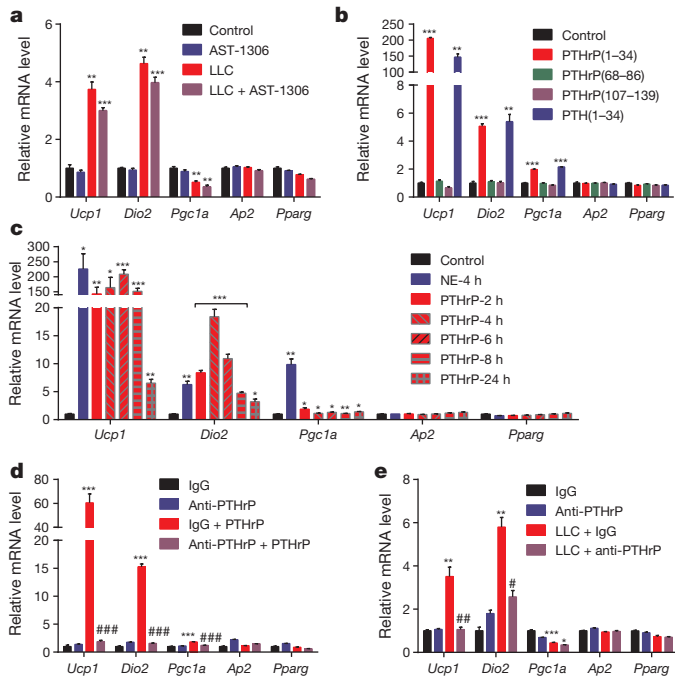


Figure 3 | PTHrP is responsible for most of the LLC-cell-derived browning activity. a–c, Primary adipocytes ($n = 3$ for each group) were treated with $1 \mu\text{M}$ AST-1306 and LLC-cell-conditioned medium for 24 h (a), with 100 ng ml^{-1} of the indicated peptides for 2 h (b), or with 100 ng ml^{-1} PTHrP for 2–24 h or 100 mM noradrenaline (NE) for 4 h (c). d, e, Primary adipocytes ($n = 3$ for each group) were treated with $10 \mu\text{g ml}^{-1}$ IgG or anti-PTHrP antibody along with either 10 ng ml^{-1} PTHrP for 4 h (d) or LLC-cell-conditioned medium for 24 h (e). The mRNA levels were measured by qRT-PCR. The values are mean \pm s.e.m. Statistical analysis was conducted using the two-tailed t -test. *, ** and *** denote differences from the control (IgG) group. #, ## and ### denote differences between the PTHrP groups or between the LLC groups. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; #, $P < 0.05$; ##, $P < 0.005$; ###, $P < 0.0005$.

PTHrP activity when added to adipocytes (Fig. 3d). Importantly, when adipocytes were treated with LLC-cell-conditioned medium in combination with the anti-PTHrP antibody, *Ucp1* and *Dio2* induction was almost completely blocked (Fig. 3e). These data indicate that PTHrP is a major component of the LLC-cell-secreted products that cause browning of cultured adipocytes.

PTHr (which is encoded by *Pth1r*) is known to be highly expressed in the kidneys and bone¹⁵. Notably, rather high *Pth1r* mRNA expression was also observed in adipose and muscle tissues (Extended Data Fig. 6a). We next investigated PTHrP-dependent transcriptional regulation *in vivo*. PTHrP potentially promoted the expression of thermogenic genes and certain other genes of energy metabolism (Extended Data Fig. 6b–e), a pattern that is notably similar to the browning of adipose tissues by LLC tumours.

The role of PTHrP in LLC-tumour-induced fat browning and wasting was then investigated. Up to 500 pg ml^{-1} PTHrP was detected in the plasma of tumour-bearing mice (Fig. 4a). We next injected mice with the neutralizing antibody specific for PTHrP or control IgG. Amazingly, anti-PTHrP-treated mice did not experience significant weight loss, while the IgG-treated group displayed evident cachexia (Fig. 4b). The anti-PTHrP treatment blocked both adipose tissue and skeletal muscle wasting, without changing the average tumour mass (Fig. 4c, d). Histological examination of the adipose tissues showed that anti-PTHrP treatment prevented the shrinkage of fat droplets (Fig. 4e). This neutralization of PTHrP also blocked thermogenic gene expression in eWAT, iBAT (Fig. 4f, g) and iWAT (Extended Data Fig. 7), indicating that thermogenesis has a causal role in fat wasting. Treatment with anti-PTHrP antibody also lowered the oxygen consumption of the cachectic mice, increased

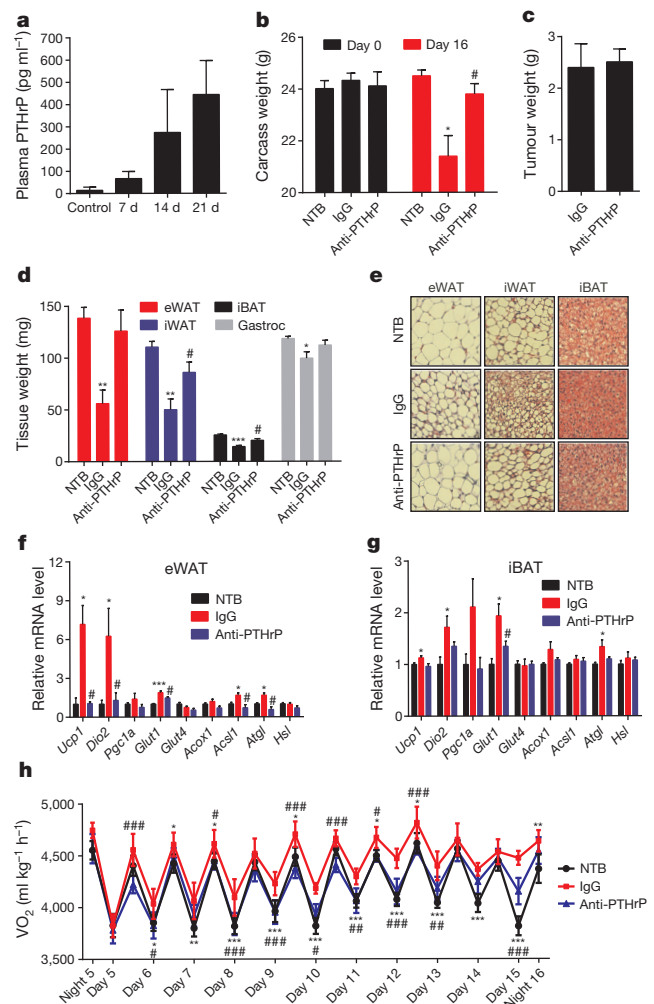


Figure 4 | Neutralization of PTHrP prevents tumour-induced adipose tissue browning. a, Plasma PTHrP concentrations in mice bearing LLC tumours up to 21 days ($n = 6$). b–g, Mice inoculated with LLC cells received 10 mg IgG or anti-PTHrP antibody per kg body weight every 3 days from day 6 to day 15 and were killed on day 16 ($n = 4, 5$ and 6 for the NTB, IgG and anti-PTHrP groups, respectively). Carcass weight (b), weight of tumours (c), weight of fat and muscle tissues (d) and haematoxylin and eosin staining of adipose tissues (e) are shown. The mRNA levels in eWAT (f) and iBAT (g) were measured by qRT-PCR. h, Mice were treated similarly to b–g except that they were placed into metabolic cages on day 4 to measure oxygen consumption ($n = 6$ for the NTB group and $n = 5$ for the other groups). The values are mean \pm s.e.m. Statistical analysis was conducted using the two-tailed t -test. *, ** and *** denote differences between the NTB group and the IgG group. #, ## and ### denote differences between the anti-PTHrP group and the IgG group. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; #, $P < 0.05$; ##, $P < 0.005$; ###, $P < 0.0005$.

their physical activity and reduced their heat production (Fig. 4h and Extended Data Fig. 8). We used a Clark electrode to test *ex vivo* oxygen consumption of fat and skeletal muscle tissues. Notwithstanding the limited power of this technique to measure and represent actual tissue respiration, we demonstrated modest increases in the fat tissue respiration of the cachectic mice. These increases were suppressed by the anti-PTHrP treatment (Extended Data Fig. 9). While additional, off-target, effects of the anti-PTHrP antibody cannot be ruled out, these data reveal a major role for PTHrP in LLC-tumour-induced adipose tissue thermogenesis and hypermetabolism.

Neutralization of PTHrP also reduced the LLC-tumour-induced muscle wasting (Fig. 4d), the atrophy of muscle fibres (Extended Data Fig. 10a) and the atrophy-associated gene expression (Fig. 5a). We further

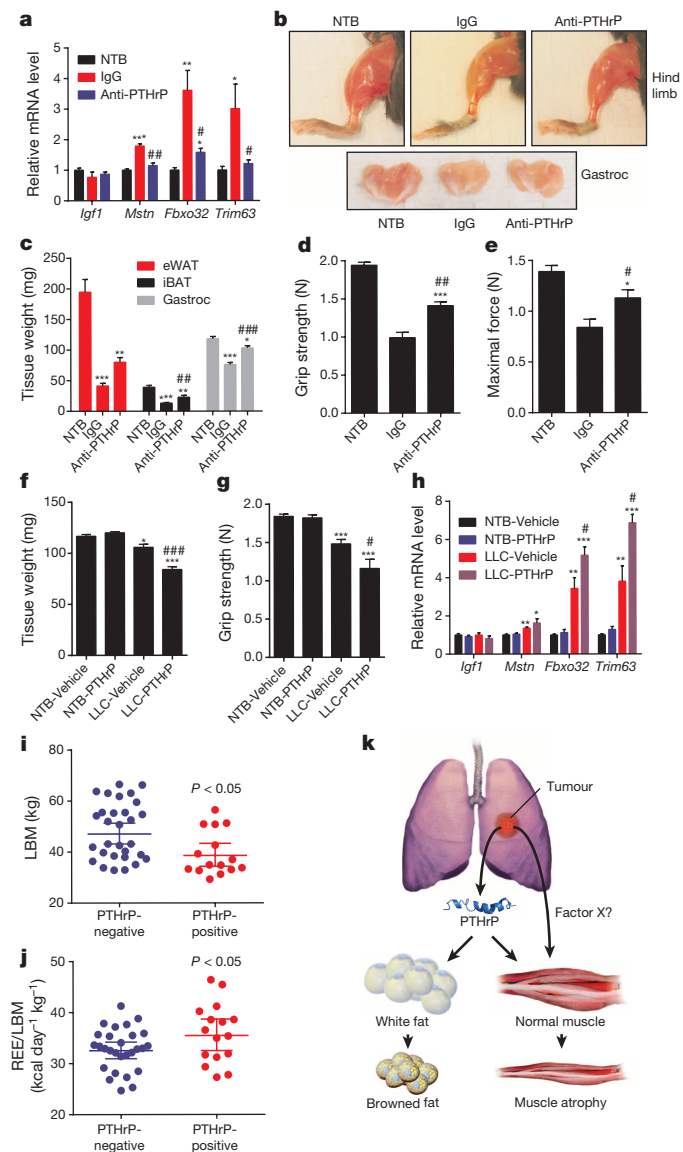


Figure 5 | PTHrP is associated with wasting of the LBM in cachectic mice and humans. **a**, The mRNA levels in the gastrocnemius muscle from the experiment described in Fig. 4b–g. **b–e**, Mice inoculated with LLC cells received 10 mg IgG or anti-PTHrP antibody per kg body weight every 3 days from day 6 to day 21–22 ($n = 5$ for the anti-PTHrP group and $n = 6$ for the other groups). Representative images of the hindlimb and gastrocnemius muscles are shown (**b**). Fat and muscle tissues were weighed (**c**). Muscle function was analysed by measuring grip strength (**d**) and *in situ* contraction (**e**). **f–h**, Mice inoculated with LLC cells were killed after receiving daily injections of 1 mg PTHrP per kg body weight between days 10 and 16 ($n = 5$ for the LLC-Vehicle group (vehicle denotes PBS) and $n = 6$ for the other groups). The gastrocnemius muscle weight (**f**) and grip strength (**g**) were measured. The mRNA levels in the gastrocnemius muscle were determined by qRT-PCR (**h**). **i, j**, The LBM and REE/LBM of the PTHrP-positive and PTHrP-negative patients were compared. **k**, A model for PTHrP action in cancer cachexia. **a, c–j**, The values are mean \pm s.e.m. Statistical analysis was conducted using the two-tailed *t*-test. *, **, and *** denote differences from the NTB group. #, ## and ### denote differences between the IgG group and the anti-PTHrP group or between the LLC-Vehicle group and the LLC-PTHrP group. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; #, $P < 0.05$; ##, $P < 0.005$; ###, $P < 0.0005$.

investigated this effect by treating mice with anti-PTHrP antibody until the time point at which severe cachexia was observed in the controls. The anti-PTHrP treatment preserved muscle mass to a significant extent (Fig. 5b, c). This preservation was also reflected in muscle function, as shown

by improved grip strength and *in situ* muscle contraction (Fig. 5d, e). While it was possible that PTHrP-associated hypercalcaemia may have confounded the effects here, we did not observe hypercalcaemia in the LLC model, and neutralization of PTHrP did not affect the circulating calcium levels (Extended Data Fig. 10b, c).

PTHrP injection into healthy mice failed to affect the expression of muscle-atrophy-associated genes (Extended Data Fig. 10d, e). Furthermore, PTHrP treatment of primary myotubes did not change the expression of the muscle-atrophy-associated genes or myotube diameter (Extended Data Fig. 10f, g), suggesting that PTHrP does not directly induce skeletal muscle atrophy. We investigated this further by injecting PTHrP into tumour-bearing mice. Remarkably, PTHrP administration exacerbated cachexia-associated weight loss without changing tumour mass (Extended Data Fig. 10h–j). PTHrP treatment led to more severe skeletal muscle wasting and poorer muscle function and significantly promoted the expression of muscle-atrophy-associated genes in the cachectic mice (Fig. 5f–h). These findings indicate that PTHrP is required for the loss of muscle mass and function in the LLC model. However, PTHrP probably does not act alone and is likely to collaborate with other tumour-derived factors.

Lastly, we investigated PTHrP expression in a cohort of patients diagnosed with metastatic non-small-cell lung cancer or colorectal cancer, whose body composition and metabolic rate were studied in the context of a clinical investigation of the aetiology of cancer cachexia^{16–18}. Seventeen out of 47 patients displayed detectable levels of serum PTHrP (mean \pm s.d., 205 ± 155 pg ml⁻¹) with no medical history of hypercalcaemia. These patients had a significantly lower lean body mass (LBM) and a significantly higher resting energy expenditure (REE) per kg LBM than 30 patients lacking detectable levels of PTHrP (Fig. 5i, j). The reduced amount of lean tissue in the PTHrP-positive patients was not associated with a difference in dietary intake (PTHrP-positive, $2,364 \pm 750$ kcal day⁻¹ (mean \pm s.d.); PTHrP-negative, $2,285 \pm 487$ kcal day⁻¹; $P = 0.66$) or a difference in systemic inflammation as determined by the serum C-reactive protein levels (PTHrP-positive, 5.7 ± 4.7 mg l⁻¹ (mean \pm s.d.); PTHrP-negative, 12.6 ± 27 mg l⁻¹; $P = 0.32$) (Supplementary Data). These results indicate an association between higher PTHrP concentrations and a greater degree of lean tissue wasting and elevated energy expenditure within this cohort of patients.

PTHrP is secreted by many tumours and is often involved in the hypercalcaemia that accompanies certain cancers¹⁹. PTHrP has previously been linked to the cachexia associated with hypercalcaemia of malignancy^{20–22}. However, its direct roles in the muscle wasting process and the browning of adipose tissues have not been reported. Taken together, our data show that PTHrP is the major LLC-tumour-derived factor stimulating adipose tissue thermogenic gene expression and hypermetabolism. Neutralization of PTHrP blocks both the browning and the wasting of fat depots and much of the muscle wasting and weakness in this model of cachexia.

These data suggest a model whereby PTHrP functions alone to drive the thermogenic/browning program in adipose tissues but is likely to collaborate with other tumour-derived molecules to bring about the profound muscle wasting and weakness observed in the LLC model (Fig. 5k). The identity of the collaborating factors is unknown. The browning driven by PTHrP seems not to be obligatorily linked to the muscle wasting process.

Our preliminary evaluation of 47 patients with cancer identified a PTHrP-positive subset, which showed features of reduced LBM and increased REE. Further prospective studies on cachexia-prone patients should study these parameters with serial tissue biopsies to determine whether there is a specific association between PTHrP, adipose tissue browning and metabolic rate in clinical cancer cachexia. If PTHrP is involved in cachexia in certain patients, it should be possible to identify those with elevated PTHrP levels. A humanized antibody specific for PTHrP might have anti-cachectic effects like those shown here. It will also be interesting to determine whether elevated PTHrP or PTH levels have any roles in the cachexia that is associated with certain other diseases, such as congestive heart failure and chronic kidney diseases.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 March; accepted 22 May 2014.

Published online 13 July 2014.

1. Bianchi, A. *et al.* Increased brown adipose tissue activity in children with malignant disease. *Horm. Metab. Res.* **21**, 640–641 (1989).
2. Bing, C. *et al.* Increased gene expression of brown fat uncoupling protein (UCP)1 and skeletal muscle UCP2 and UCP3 in MAC16-induced cancer cachexia. *Cancer Res.* **60**, 2405–2410 (2000).
3. Brooks, S. L., Neville, A. M., Rothwell, N. J., Stock, M. J. & Wilson, S. Sympathetic activation of brown-adipose-tissue thermogenesis in cachexia. *Biosci. Rep.* **1**, 509–517 (1981).
4. Roe, S., Cooper, A. L., Morris, I. D. & Rothwell, N. J. Mechanisms of cachexia induced by T-cell leukemia in the rat. *Metabolism* **45**, 645–651 (1996).
5. Shellock, F. G., Riedinger, M. S. & Fishbein, M. C. Brown adipose tissue in cancer patients: possible cause of cancer-induced cachexia. *J. Cancer Res. Clin. Oncol.* **111**, 82–85 (1986).
6. Tsoli, M. *et al.* Activation of thermogenesis in brown adipose tissue and dysregulated lipid metabolism associated with cancer cachexia in mice. *Cancer Res.* **72**, 4372–4382 (2012).
7. Fearon, K. C., Glass, D. J. & Guttridge, D. C. Cancer cachexia: mediators, signaling, and metabolic pathways. *Cell Metab.* **16**, 153–166 (2012).
8. Ovesen, L., Allingstrup, L., Hannibal, J., Mortensen, E. L. & Hansen, O. P. Effect of dietary counseling on food intake, body weight, response rate, survival, and quality of life in cancer patients undergoing chemotherapy: a prospective, randomized study. *J. Clin. Oncol.* **11**, 2043–2049 (1993).
9. Tisdale, M. J. Mechanisms of cancer cachexia. *Physiol. Rev.* **89**, 381–410 (2009).
10. Penna, F. *et al.* Anti-cytokine strategies for the treatment of cancer-related anorexia and cachexia. *Expert Opin. Biol. Ther.* **10**, 1241–1250 (2010).
11. Fearon, K., Arends, J. & Baracos, V. Understanding the mechanisms and treatment options in cancer cachexia. *Nature Rev. Clin. Oncol.* **10**, 90–99 (2013).
12. Wu, J. *et al.* Beige adipocytes are a distinct type of thermogenic fat cell in mouse and human. *Cell* **150**, 366–376 (2012).
13. Cohen, P. *et al.* Ablation of PRDM16 and beige adipose causes metabolic dysfunction and a subcutaneous to visceral fat switch. *Cell* **156**, 304–316 (2014).
14. Das, S. K. *et al.* Adipose triglyceride lipase contributes to cancer-associated cachexia. *Science* **333**, 233–238 (2011).
15. Vilardaga, J. P., Romero, G., Friedman, P. A. & Gardella, T. J. Molecular basis of parathyroid hormone receptor signaling and trafficking: a family B GPCR paradigm. *Cell. Mol. Life Sci.* **68**, 1–13 (2011).
16. Liefers, J. R. *et al.* A viscerally driven cachexia syndrome in patients with advanced colorectal cancer: contributions of organ and tumor mass to whole-body energy demands. *Am. J. Clin. Nutr.* **89**, 1173–1179 (2009).
17. Mourtzakis, M. *et al.* A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl. Physiol. Nutr. Metab.* **33**, 997–1006 (2008).
18. Prado, C. M. *et al.* Dietary patterns of patients with advanced lung or colorectal cancer. *Can. J. Diet. Pract. Res.* **73**, e298–e303 (2012).
19. Mundy, G. R. & Edwards, J. R. PTH-related peptide (PTHrP) in hypercalcemia. *J. Am. Soc. Nephrol.* **19**, 672–675 (2008).
20. Iguchi, H., Aramaki, Y., Maruta, S. & Takiguchi, S. Effects of anti-parathyroid hormone-related protein monoclonal antibody and osteoprotegerin on PTHrP-producing tumor-induced cachexia in nude mice. *J. Bone Miner. Metab.* **24**, 16–19 (2006).
21. Iguchi, H., Onuma, E., Sato, K., Sato, K. & Ogata, E. Involvement of parathyroid hormone-related protein in experimental cachexia induced by a human lung cancer-derived cell line established from a bone metastasis specimen. *Int. J. Cancer* **94**, 24–27 (2001).
22. Sato, K. *et al.* Passive immunization with anti-parathyroid hormone-related protein monoclonal antibody markedly prolongs survival time of hypercalcemic nude mice bearing transplanted human PTHrP-producing tumors. *J. Bone Miner. Res.* **8**, 849–860 (1993).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Mourtzakis and C. Prado for their help with the human study. S. Kir is a Robert Black Fellow of the Damon Runyon Cancer Research Foundation (DRG-2153-13), and J.P.W. is supported by a postdoctoral fellowship from the American Cancer Society (PF-13-385-01-TBE). This work was supported by National Institutes of Health grant DK31405 to B.M.S.

Author Contributions S. Kir and B.M.S. conceived and designed the experiments. S. Kir, J.P.W., S. Kleiner, L.K. and P.C. performed the experiments. S. Kir, J.P.W. and V.E.B. analysed the data. S. Kir and B.M.S. wrote the manuscript.

Author Information The microarray data set has been deposited in the Gene Expression Omnibus (GEO) database under the accession GSE57797. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.M.S. (bruce_spiegelman@dfci.harvard.edu).

eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies

Lise Boussemart^{1,2,3*}, Hélène Malka-Mahieu^{1,2*}, Isabelle Girault^{1*}, Delphine Allard¹, Oskar Hemmingsson^{1†}, Gorana Tomasic⁴, Marina Thomas³, Christine Basmadjian⁵, Nigel Ribeiro⁵, Frédéric Thuaud⁵, Christina Mateus³, Emilie Routier³, Nyam Kamsu-Kom¹, Sandrine Agoussi¹, Alexander M. Eggermont^{2,3}, Laurent Désaubry⁵, Caroline Robert^{1,2,3} & Stéphan Vagner^{1,2,3†}

In BRAF(V600)-mutant tumours, most mechanisms of resistance to drugs that target the BRAF and/or MEK kinases rely on reactivation of the RAS–RAF–MEK–ERK mitogen-activated protein kinase (MAPK) signal transduction pathway, on activation of the alternative, PI(3)K–AKT–mTOR, pathway (which is ERK independent) or on modulation of the caspase-dependent apoptotic cascade^{1–3}. All three pathways converge to regulate the formation of the eIF4F eukaryotic translation initiation complex, which binds to the 7-methylguanylate cap (m⁷G) at the 5' end of messenger RNA, thereby modulating the translation of specific mRNAs^{4,5}. Here we show that the persistent formation of the eIF4F complex, comprising the eIF4E cap-binding protein, the eIF4G scaffolding protein and the eIF4A RNA helicase, is associated with resistance to anti-BRAF, anti-MEK and anti-BRAF plus anti-MEK drug combinations in BRAF(V600)-mutant melanoma, colon and thyroid cancer cell lines. Resistance to treatment and maintenance of eIF4F complex formation is associated with one of three mechanisms: reactivation of MAPK signalling, persistent ERK-independent phosphorylation of the inhibitory eIF4E-binding protein 4EBP1 or increased pro-apoptotic BCL-2-modifying factor (BMF)-dependent degradation of eIF4G. The development of an *in situ* method to detect the eIF4E–eIF4G interactions shows that eIF4F complex formation is decreased in tumours that respond to anti-BRAF therapy and increased in resistant metastases compared to tumours before treatment. Strikingly, inhibiting the eIF4F complex, either by blocking the eIF4E–eIF4G interaction or by targeting eIF4A, synergizes with inhibiting BRAF(V600) to kill the cancer cells. eIF4F not only appears to be an indicator of both innate and acquired resistance but also is a promising therapeutic target. Combinations of drugs targeting BRAF (and/or MEK) and eIF4F may overcome most of the resistance mechanisms arising in BRAF(V600)-mutant cancers.

A plethora of mechanisms of resistance to BRAF(V600) inhibitors have been described, including mutations of *NRAS*, *MAP2K1* and *PIK3CA*, partial deletion of *PTEN*, expression of *BRAF* splice variants, upregulation of PDGFR β , IGF-1R, EGFR, HER3 and COT and downregulation of BH3-only proteins (which are pro-apoptotic)^{1–3,6}. By activating the MAPK or the PI(3)K pathway or by modulating the caspase-dependent apoptotic cascade, all of these mechanisms converge on eIF4F-dependent mRNA translation.

To investigate the potential role of eIF4F complex formation in the resistance of BRAF(V600) melanoma to anti-BRAF and anti-MEK compounds, we first tested a panel of human melanoma cell lines (Supplementary Table 1) for their sensitivity to vemurafenib. Two BRAF(V600)-mutated cell lines (Mel624 and A2058) were resistant to vemurafenib compared with other cell lines (for example, A375 and Mel888) in both short-term (Extended Data Fig. 1a) and long-term (Extended Data Fig. 1b) proliferation assays. Similarly to the sensitivity pattern observed *in vitro*,

A375 xenografts were more sensitive to vemurafenib than were Mel624 xenografts (Extended Data Fig. 2a). Determination of the half-maximum inhibitory concentration (IC₅₀) (Extended Data Fig. 1a and Supplementary Table 1), together with long-term proliferation assays (Extended Data Fig. 1c), indicated that these two vemurafenib-resistant cell lines (Mel624 and A2058) were also less sensitive to dabrafenib (another BRAF(V600) inhibitor), PD0325901 and trametinib (both MEK inhibitors), the combination of dabrafenib and trametinib, and CGP 57380 (a MNK inhibitor). To determine the status of the eIF4F complex in these various cell lines, we used a cap-binding assay in which synthetic m⁷GTP-Sepharose beads capture eIF4E and its two binding partners, eIF4G and 4EBP1. We observed that vemurafenib treatment increased the amount of eIF4E-bound 4EBP1 and concomitantly decreased the amount of eIF4E-bound eIF4G in the two sensitive cell lines (A375 and Mel888) (Fig. 1a). Incorporation of [³⁵S]methionine and cysteine and profiling of polysomes showed that this decrease in eIF4F complex formation led to a weak decrease (13%) in translation initiation in the acute phase (3 h) of the response to treatment and to a greater decrease (30%) in protein synthesis after 24 h of treatment (Extended Data Fig. 3a, b). As expected from another study⁷, describing mTOR-dependent translation regulation, the analysis of mRNAs that are differentially recruited to heavy polysomes (Supplementary Table 2) showed that the translation of several mRNAs containing 5' terminal oligopyrimidine (TOP) motifs and encoding many ribosomal proteins and translation factors was inhibited during the acute response to vemurafenib treatment (Extended Data Figs 3c, d and 4 and Supplementary Table 2).

Strikingly, the vemurafenib-dependent disruption of the eIF4E–eIF4G complex was not observed in the resistant cell lines (Mel624 and A2058) (Fig. 1a). We further investigated the formation of the eIF4F complex using a proximity ligation assay procedure (Extended Data Fig. 5), and we observed that vemurafenib induced a decrease in eIF4E–eIF4G interaction and an increase in eIF4E–4EBP1 complex formation in vemurafenib-sensitive cell lines (A375 and Mel888) that was not observed in the resistant cell lines (Mel624 and A2058) (Fig. 1b and Extended data Fig. 6a, b). *In vivo* experiments in mice were consistent with these findings, as we also found fewer eIF4E–eIF4G interactions, compared with eIF4E–4EBP1 interactions, in vemurafenib-treated A375 xenografts (sensitive) than in vemurafenib-treated Mel624 xenografts (resistant) (Extended Data Fig. 2b, c). The same correlation between sensitivity to the drugs and the level of activation of the eIF4F complex was observed with dabrafenib, trametinib and combinations of both drugs (Extended Data Fig. 7a, b), extending our observation to anti-MEK agents and anti-BRAF plus anti-MEK combinations⁸.

In parallel, we generated a melanoma cell line with acquired resistance to vemurafenib (Supplementary Table 1) by exposing the Malme-3M cell line, which has the BRAF(V600) mutation, to incremental increases

¹Inserm UMR981, Villejuif F-94805, France. ²Université Paris-Sud XI, Kremlin-Bicêtre F-94276, France. ³Gustave Roussy, Dermato-Oncology, Villejuif F-94805, France. ⁴Gustave Roussy, Pathology Department, Villejuif F-94805, France. ⁵CNRS-Strasbourg University, UMR7200, Illkirch F-67400, France. [†]Present addresses: Department of Surgical and Perioperative Sciences, Umeå University, Umeå SE-90187, Sweden (O.H.); CNRS UMR3348, Institut Curie, Orsay F-91405, France (S.V.).

*These authors contributed equally to this work.

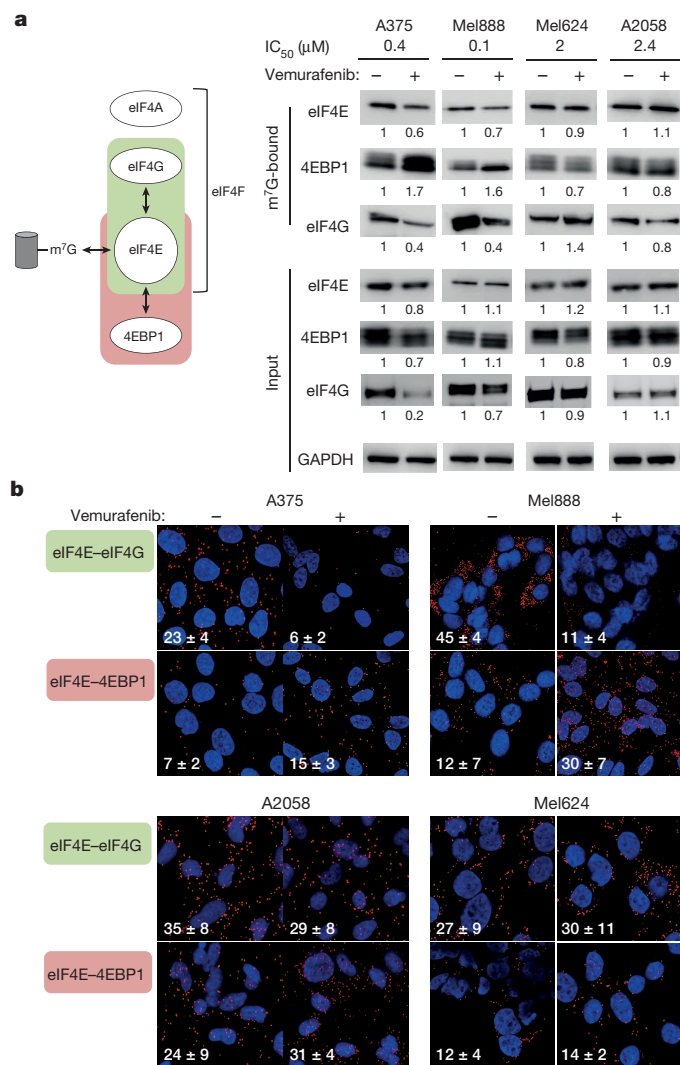


Figure 1 | The formation of the eIF4F translation initiation complex is associated with resistance to BRAF inhibitors. **a**, eIF4F complex formation, as determined using the cap-binding assay in vemurafenib-treated or untreated tumour cell lines. Eluates from beads (m⁷G-bound) and total extracts (input) were analysed by western blotting. The half-maximum inhibitory concentration (IC₅₀) for vemurafenib is indicated (see also Supplementary Table 1). **b**, eIF4E-eIF4G or eIF4E-4EBP1 interactions detected by proximity ligation assay (PLA) in vemurafenib-treated or untreated cell lines. The interactions were visualized as red spots. The white numbers indicate the mean ± s.d. (*n* = 3).

in vemurafenib concentration over 2 months. Again, vemurafenib induced a decrease in eIF4E-eIF4G interactions and an increase in eIF4E-4EBP1 complex formation in vemurafenib-sensitive Malme-3M cells but not in vemurafenib-resistant Malme-3M (R-Malme-3M) cells (Extended Data Fig. 8a). Finally, this paradigm was not limited to melanoma cell lines: we observed the maintenance of the eIF4F complex in two other cell lines: the colon cancer cell line HT-29, in which resistance is associated with an activating PI(3)K(P449T) mutation⁹ and *EGFR* overexpression¹⁰; and the papillary thyroid cancer cell line BCPAP, in which resistance is linked to *HER3* overexpression¹¹ (Extended Data Fig. 8b).

We next addressed the mechanisms underlying the differences in the sensitivity to vemurafenib in the above-mentioned melanoma cell lines. We observed that R-Malme-3M exhibited a higher ERK phosphorylation level than the sensitive Malme-3M cell line (Fig. 2a). The sustained phosphorylation of ERK was associated with a vemurafenib-dependent increase in BRAF-CRAF dimerization, revealing a mechanism of resistance linked to reactivation of the MAPK pathway (Fig. 2b). Vemurafenib

treatment inhibited ERK phosphorylation in all of the other cell lines tested, including the vemurafenib-resistant cell lines Mel624 and A2058 (Fig. 2c), showing that the differential sensitivity to vemurafenib in these cell lines cannot be explained by the reactivation of the MAPK pathway. The expression and phosphorylation of the main translation initiation factors remained unchanged in vemurafenib-resistant cell lines, except for eIF4G and eIF4B, for which vemurafenib-dependent dephosphorylation was observed (Fig. 2c). However, these dephosphorylations also occurred in vemurafenib-sensitive cell lines and therefore were not correlated with vemurafenib resistance. 4EBP1 was dephosphorylated in the vemurafenib-treated sensitive cell line Mel888 and to a lesser extent in A375. This dephosphorylation is linked to AKT dephosphorylation (Fig. 2c), which does not occur in resistant cell lines. The amounts of eIF4G and eIF4B were decreased in the vemurafenib-treated sensitive A375 cell line (Fig. 2c).

To uncover the mechanisms underlying this decrease, we sought to identify the genes whose expression is altered after vemurafenib treatment in A375 cells compared with Mel624 cells, by using microarrays (Supplementary Table 3). As expected from previous studies^{12,13}, vemurafenib led to the inhibition of expression of genes encoding transcription factors associated with ERK-dependent transformation and genes encoding feedback regulators of ERK signalling, such as members of the dual specificity phosphatase (DUSP) and sprouty (SPRY) gene families (Fig. 2d and Supplementary Table 3). These changes were found in both the sensitive and resistant cell lines, reinforcing the finding that the differential sensitivity of these two cell lines to vemurafenib is independent of the MAPK pathway. We focused on a set of 27 genes that were differentially upregulated in A375 cells, including *BMF* (Fig. 2d), a pro-apoptotic gene that had already been shown to be involved in acquired resistance to PLX4720 (ref. 6), a vemurafenib analogue. Of these 27 genes, *BMF* was the only one that, when knocked down (by short interfering RNA (siRNA)-mediated depletion), led to the relative resistance of A375 cells to vemurafenib (Extended Data Fig. 9a, b). Vemurafenib treatment led to a strong increase in the expression of *BMF* in A375 cells but not in Mel624 cells (Fig. 2e). siRNA-mediated depletion of *BMF* led to decreased vemurafenib-dependent eIF4G cleavage (Fig. 2f) and to increased eIF4F complex formation (Fig. 2g), suggesting that *BMF* is involved in the sensitivity to vemurafenib by acting on the cleavage of eIF4G, which, in turn, affects eIF4F complex formation. In favour of this hypothesis, we found that the cleavage of eIF4G (first observed 45 min after treatment) preceded the disruption of eIF4F (first observed 90 min after treatment) (Extended Data Fig. 6c, d). Thus, in addition to reactivation of the MAPK pathway and/or activation of the PI(3)K-AKT-mTOR pathway, the formation of the eIF4F complex is also an indicator of resistance linked to modulation of apoptotic cascades.

To further support the idea that the formation of the eIF4F complex is involved in sensitivity to vemurafenib, we used the small molecule inhibitor 4EGI-1 (4EGI-1), which disrupts eIF4E-eIF4G interaction¹⁴. 4EGI-1 decreased the eIF4E-eIF4G interaction and, in parallel, increased the eIF4E-4EBP1 interaction in all cell lines tested (Extended Data Fig. 5b). Vemurafenib-resistant cell lines were more sensitive to 4EGI-1 than were vemurafenib-sensitive cell lines (Fig. 3a), suggesting a direct role of the eIF4F complex in proliferation. To examine whether resistance can be reversed by disrupting the eIF4F complex, we explored the effect of the combination of vemurafenib and 4EGI-1 on Mel624 cells and found a clear synergistic effect of these compounds (Fig. 3b). Thus, the formation of the eIF4F complex is directly involved in sensitivity to vemurafenib. Targeting translation initiation may therefore be a strategy to overcome resistance to anti-BRAF and/or anti-MEK compounds. Silvestrol, a naturally occurring member of the flavagline family of compounds, inhibits cap-dependent translation by targeting eIF4A¹⁵⁻¹⁷. However, silvestrol displays poor ADME (absorption, distribution, metabolism and excretion) characteristics and is sensitive to P-glycoprotein-mediated multidrug resistance, which severely hampers its use in *in vivo* studies^{18,19}.

To circumvent this problem, we chose to evaluate a series of synthetic flavagline derivatives (Supplementary Notes) for their capacity

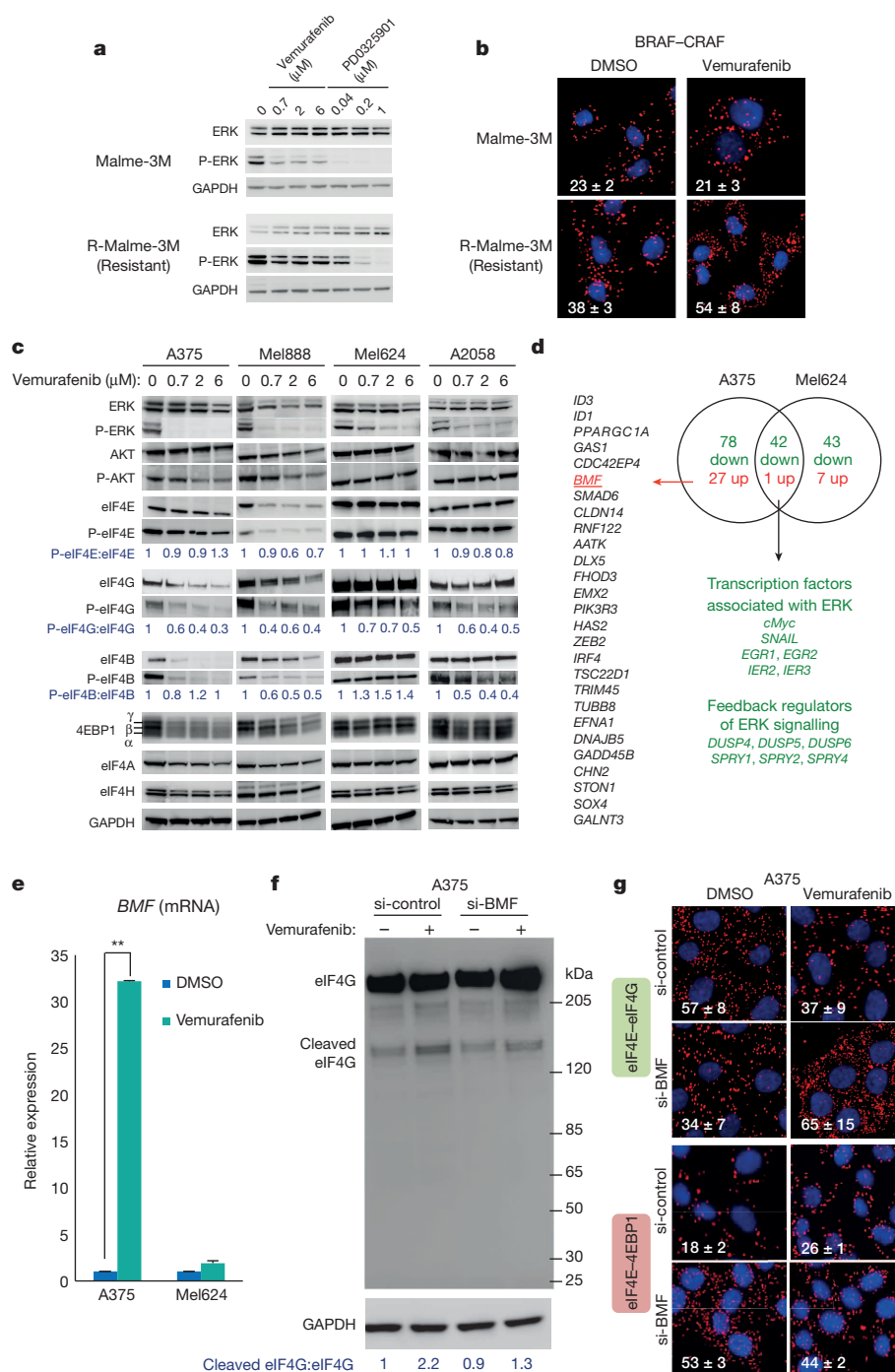


Figure 2 | Identification of mechanisms of resistance associated with persistent eIF4F formation. **a, c**, Western blot analysis with antibodies specific for phosphorylated (P-) or total proteins as indicated. **b**, BRAF-CRAF interaction detected by PLA. The white numbers indicate the mean \pm s.d. ($n = 3$). **d**, Venn diagram showing the overlap of differentially expressed genes in several vemurafenib-treated cell lines. The list of genes on the left represents the 27 genes whose expression was upregulated by vemurafenib in A375 cells but not in Mel624 cells. *EGR*, early growth response; *IER*, immediate early

response. **e**, Expression of *BMF* mRNA (as determined by quantitative reverse transcription PCR (qRT-PCR)) in vemurafenib-treated A375 cells compared with dimethylsulphoxide (DMSO) controls. The data are presented as the mean \pm s.d. ($n = 3$), and differences were assessed with Student's *t*-test (**, $P < 0.01$). **f**, Western blotting analysis of vemurafenib treatment on *BMF*-depleted A375 cells. si-*BMF*, *BMF*-directed siRNA; si-control, control siRNA. **g**, eIF4E-eIF4G and eIF4E-4EBP1 interactions detected by PLA. The white numbers are the mean \pm s.d. ($n = 3$).

to inhibit the eIF4F complex, by using a bicistronic luciferase reporter construct (pR-HepC-L) (Fig. 3c). In this assay, translation of the first (LucR) cistron is eIF4F-dependent, while translation of the hepatitis C virus internal ribosome entry site (IRES)-driven second (LucF) cistron is eIF4F-independent. Differential translation of both cistrons, as measured by the ratio of luciferase activities, directly reflects differences in eIF4F activity. Treatment of pR-HepC-L-transfected A375 cells with

silvestrol and several flavaglines (for example, FL3, FL14, FL23 or FL36) led to an increase in the LucF:LucR ratio (Fig. 3d), indicating that those drugs target eIF4F in a specific manner. A structure-function relationship study allowed us to define the best translation inhibitors (Fig. 3e). The ability to inhibit eIF4F-dependent translation correlated with the ability to inhibit cell proliferation (Fig. 3f). Finally, we found that combining vemurafenib and an eIF4A inhibitor, such as FL3, silvestrol or

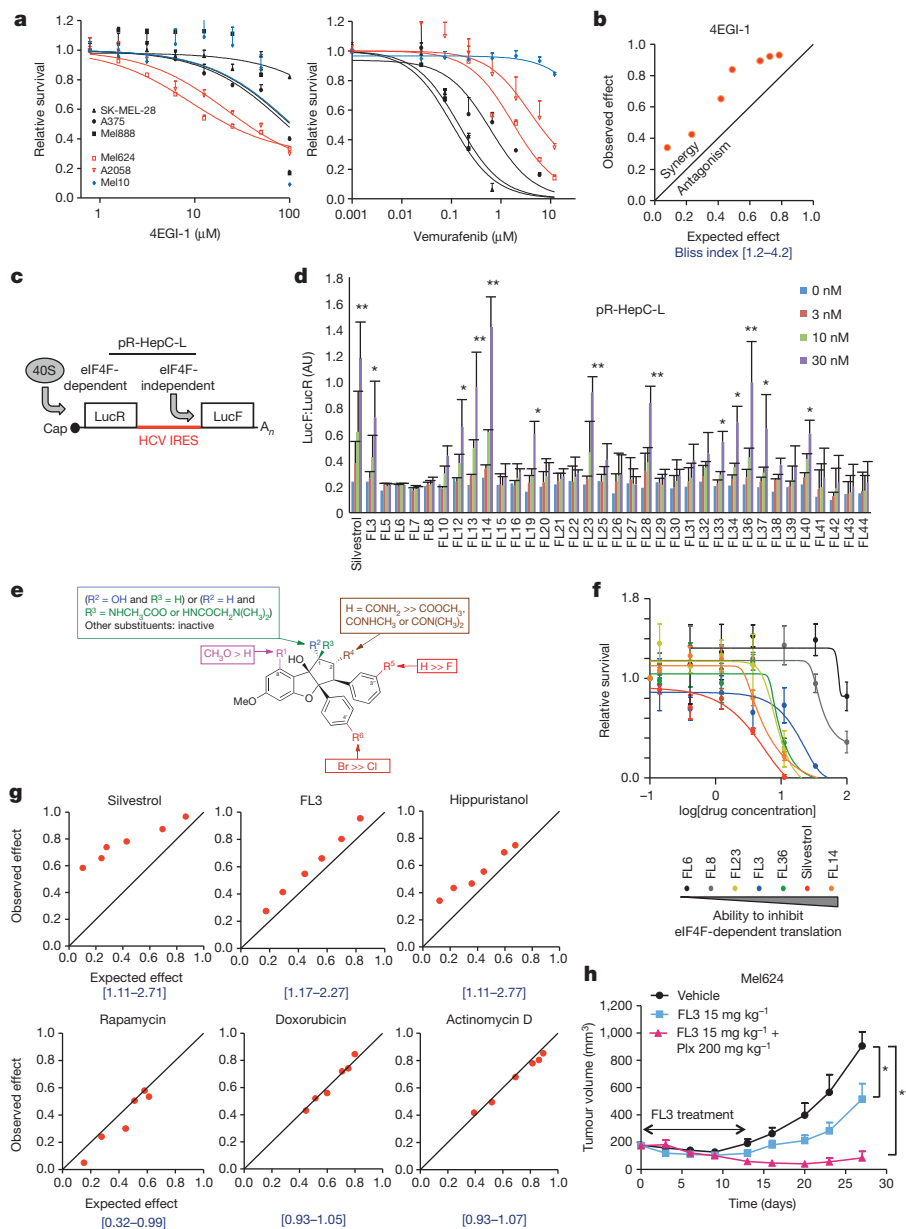


Figure 3 | Flavaglines target eIF4F and synergize with vemurafenib. **a, f**, Cell proliferation assays. The data are presented as the mean \pm s.d. ($n = 4$). Survival is shown relative to DMSO controls. (The ability to inhibit eIF4F-dependent translation is derived from **d**.) **b, g**, Isobologram of the effect of the combination of vemurafenib plus the indicated drugs on Mel624 cells. The Bliss index is shown in square brackets. **c**, Schematic of the mRNA produced from the pR-HepC-L vector. HCV IRES, hepatitis C virus internal ribosome entry site. **d**, Ratio of LucF to LucR activity (AU, arbitrary units) in pR-HepC-L-transfected A375 cells treated with flavaglines. The data are presented as the mean \pm s.d. ($n = 4$), and differences were assessed with Student's *t*-test (*, $P < 0.05$; **, $P < 0.01$). **e**, Structure–function relationships of the flavaglines. **h**, Growth of Mel624 xenografts in mice. The data are presented as mean tumour volume \pm s.e.m. ($n = 6$ mice per group); one-sided Mann–Whitney test, *, $P < 0.05$; **, $P < 0.01$.

hippuristanol⁸, had a synergistic effect on the proliferation of Mel624 cells (Fig. 3g). Interestingly, this synergy was not observed with other chemotherapeutics such as doxorubicin (which intercalates into duplex DNA), actinomycin D (which inhibits RNA synthesis) or rapamycin (which targets mTORC1) (Fig. 3g). Rapamycin is expected to dephosphorylate 4EBP1 and should disrupt the eIF4F complex, which, in turn, should lead to a synergistic effect of vemurafenib and rapamycin. However, we found that rapamycin did not lead to 4EBP1 dephosphorylation in Mel624 cells, explaining the lack of synergy (Extended Data Fig. 9c). Xenograft experiments were then performed with FL3, which was chosen from the flavaglines because of its ability to overcome multidrug resistance and its activity *in vivo*²⁰. The combination of FL3 plus PLX4720 (Plx) given at a moderate dose of 200 mg per kg body weight led to a strong reduction in Mel624 tumour growth (Fig. 3h). The cooperative action of FL3 and Plx was shown by using a low dose of Plx (90 mg per kg body weight), at which no effect on Mel624 tumour growth was observed with Plx alone (Extended Data Fig. 2d).

To evaluate the clinical relevance of the eIF4F complex in patients, we investigated its presence in 16 biopsy specimens (before and during anti-BRAF therapy) from 7 patients with metastatic melanoma (Supplementary

Table 4). Interestingly, sequencing of the tumours that were progressing revealed an acquired PI(3)K(Q546R) mutation in a metastasis from patient 6 (Supplementary Table 5) close to a recently described activating mutation (E545G) in PI(3)K that is associated with acquired resistance². We observed a decrease in eIF4F complex formation (as measured by the ratio of the eIF4E–eIF4G complex to the eIF4E–4EBP1 complex) in three out of four biopsy samples of responding tumours (Fig. 4 and Extended Data Fig. 10) compared with biopsy samples taken before therapy (baseline). By contrast, in all resistant tumours that were progressing under therapy, we found a clear increase in eIF4F complex formation compared with the baseline biopsy samples (Fig. 4 and Extended Data Fig. 10). Interestingly, patient 4 had two metastases with an early mixed response to vemurafenib treatment: one metastasis responding to treatment (meta. 1), and the other resisting treatment (meta. 2). While an increase in eIF4F was found in the resistant metastasis, no change was observed in the responding metastasis (Fig. 4b). Thus, a direct correlation was found between the response to treatment and the formation of eIF4F complexes, illustrating the *in vivo* association between sensitivity to anti-BRAF therapy and inhibition of the eIF4F complex. We found that the increase in eIF4F complex formation was

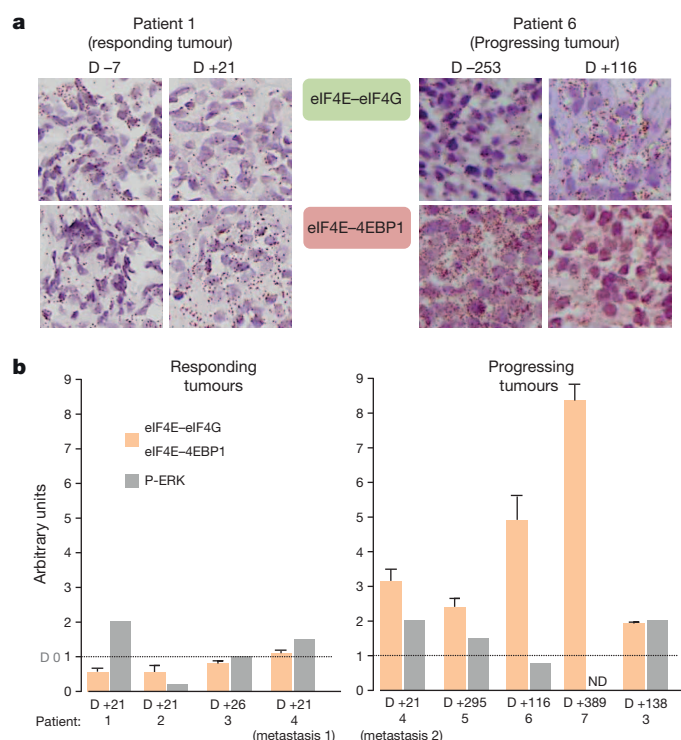


Figure 4 | Formation of the eIF4F complex in patients. **a**, eIF4E-eIF4G and eIF4E-4EBP1 interactions detected by PLA in samples from the indicated patients at the indicated times. The number corresponds to the day (D) of biopsy before (–) or after (+) treatment. Each brown spot represents an interaction. **b**, PLA quantification showing the ratio between the eIF4E-eIF4G interactions per 100 nuclei and the eIF4E-4EBP1 interactions per 100 nuclei, normalized to the same ratio before treatment (D 0) (mean \pm s.d.). Phosphorylated ERK (P-ERK) immunoreactivity was assessed semiquantitatively (grey bars).

more significantly associated with resistance than was the ERK phosphorylation level, which is conventionally investigated to explore resistance (Fig. 4b). This finding confirmed our *in vitro* data showing that the formation of the eIF4F complex was associated with resistance to vemurafenib in both ERK-dependent and ERK-independent resistance mechanisms. Using emerging technologies, such as ribosome profiling²¹, to examine the translation landscape of anti-BRAF and/or anti-MEK-sensitive and -resistant tumour cell lines should pave the way for an in-depth understanding of the contribution of genes that are regulated at the translational level in resistance mechanisms to targeted cancer therapies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 December 2013; accepted 10 June 2014.

Published online 27 July 2014.

1. Lito, P., Rosen, N. & Solit, D. B. Tumour adaptation and resistance to RAF inhibitors. *Nature Med.* **19**, 1401–1409 (2013).
2. Shi, H. *et al.* Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. *Cancer Discov.* **4**, 80 (2014).
3. Tentori, L., Lacal, P. M. & Graziani, G. Challenging resistance mechanisms to therapies for metastatic melanoma. *Trends Pharmacol. Sci.* **34**, 656–666 (2013).

4. Blagden, S. P. & Willis, A. E. The biological and therapeutic relevance of mRNA translation in cancer. *Nature Rev. Clin. Oncol.* **8**, 280–291 (2011).
5. Silvera, D., Formenti, S. C. & Schneider, R. J. Translational control in cancer. *Nature Rev. Cancer* **10**, 254–266 (2010).
6. Shao, Y. & Aplin, A. E. BH3-only protein silencing contributes to acquired resistance to PLX4720 in human melanoma. *Cell Death Differ.* **19**, 2029–2039 (2012).
7. Thoreen, C. C. *et al.* A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485**, 109–113 (2012).
8. Cencic, R., Galicia-Vazquez, G. & Pelletier, J. Inhibitors of translation targeting eukaryotic translation initiation factor 4A. *Methods Enzymol.* **511**, 437–461 (2012).
9. Oikonomou, E., Koc, M., Sourkova, V., Andera, L. & Pintzas, A. Selective BRAFV600E inhibitor PLX4720, requires TRAIL assistance to overcome oncogenic PIK3CA resistance. *PLoS ONE* **6**, e21632 (2011).
10. Prahalad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* **483**, 100–103 (2012).
11. Montero-Conde, C. *et al.* Relief of feedback inhibition of *HER3* transcription by RAF and MEK inhibitors attenuates their antitumor effects in *BRAF*-mutant thyroid carcinomas. *Cancer Discov.* **3**, 520–533 (2013).
12. Packer, L. M., East, P., Reis-Filho, J. S. & Marais, R. Identification of direct transcriptional targets of *V600E*BRAF/MEK signalling in melanoma. *Pigment Cell Melanoma Res.* **22**, 785–798 (2009).
13. Pratilas, C. A. *et al.* *V600E*BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. *Proc. Natl Acad. Sci. USA* **106**, 4519–4524 (2009).
14. Moerke, N. J. *et al.* Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128**, 257–267 (2007).
15. Bordeleau, M. E. *et al.* Therapeutic suppression of translation initiation modulates chemosensitivity in a mouse lymphoma model. *J. Clin. Invest.* **118**, 2651–2660 (2008).
16. Cencic, R. *et al.* Synergistic effect of inhibiting translation initiation in combination with cytotoxic agents in acute myelogenous leukemia cells. *Leuk. Res.* **34**, 535–541 (2010).
17. Sadlish, H. *et al.* Evidence for a functionally relevant rocamide binding site on the eIF4A-RNA complex. *ACS Chem. Biol.* **8**, 1519–1527 (2013).
18. Gupta, S. V. *et al.* Resistance to the translation initiation inhibitor silvestrol is mediated by ABCB1/P-glycoprotein overexpression in acute lymphoblastic leukemia cells. *AAPS J.* **13**, 357–364 (2011).
19. Liu, T. *et al.* Synthetic silvestrol analogues as potent and selective protein synthesis inhibitors. *J. Med. Chem.* **55**, 8859–8878 (2012).
20. Thuaud, F. *et al.* Synthetic analogue of rocamide displays a potent and selective cytotoxicity in cancer cells: involvement of apoptosis inducing factor and caspase-12. *J. Med. Chem.* **52**, 5176–5187 (2009).
21. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols* **7**, 1534–1550 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors thank J. Tanaka for providing hippuristanol. We thank the following Gustave Roussy platforms: Imaging and Cytometry Platform IRCIV (S. Salome-Desmoulez), Module de Développement en Pathologie, SIRIC SOCRATE (J. Adam), Translational Research Laboratory and Biobank (M. Breckler and L. Lacroix), Plateforme d'évaluation Préclinique (P. Gonin and K. Ser-le Roux), Genomic Core Facility (N. Pata-Merci) and Bioinformatic Core Facility (G. Meurice). We also thank V. Camara-Clayette for help with ³⁵S experiments, S. Roy for patient data collection and L. Saint Ange for text editing. C.R. and S.V.'s team was supported by Institut National du Cancer (INCA), Association pour la Recherche sur le Cancer (ARC) and Ligue contre le Cancer via an Integrated Research Action Program Melanoma (PAIR Melanome), Cancéropôle Ile de France and Ensemble Contre le Mélanome. L.D. was supported by the Association pour la Recherche sur le Cancer (ARC). We also thank the ARC and AAREC Filia Research for fellowships to N.R. and C.B. O.H. was supported by the Wenner-Gren Foundation and the Swedish Society of Medicine.

Author Contributions L.B., H.M.-M., I.G., D.A., O.H., G.T., C.B., N.R., F.T., N.K.-K., S.A. and L.D. designed and performed experiments and analysed the data. C.M., E.R., M.T. and A.M.E. provided clinical samples and gave advice. S.V. and C.R. supervised all research, wrote the manuscript and are joint senior authors.

Author Information Microarray data have been deposited in the ArrayExpress database under accession number E-MTAB-2607. Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of the paper. The authors declare competing financial interests: details are available in the online version of the paper. Correspondence and requests for materials should be addressed to C.R. (caroline.robert@gustaveroussy.fr) or S.V. (stephan.vagner@inserm.fr).

Mutant IDH inhibits HNF-4 α to block hepatocyte differentiation and promote biliary cancer

Supriya K. Saha^{1*}, Christine A. Parachoniak^{1*}, Krishna S. Ghanta¹, Julien Fitamant¹, Kenneth N. Ross¹, Mortada S. Najem¹, Sushma Gurumurthy¹, Esra A. Akbay², Daniela Sia^{3,4,5}, Helena Cornella³, Oriana Miltiadous⁴, Chad Walesky⁶, Vikram Deshpande¹, Andrew X. Zhu¹, Aram F. Hezel⁷, Katharine E. Yen⁸, Kimberly S. Straley⁸, Jeremy Travins⁸, Janeta Popovici-Muller⁸, Camelia Gliser⁸, Cristina R. Ferrone¹, Udayan Apte⁶, Josep M. Llovet^{3,4,9,10}, Kwok-Kin Wong², Sridhar Ramaswamy^{1,11} & Nabeel Bardeesy¹

Mutations in isocitrate dehydrogenase 1 (*IDH1*) and *IDH2* are among the most common genetic alterations in intrahepatic cholangiocarcinoma (IHCC), a deadly liver cancer^{1–5}. Mutant IDH proteins in IHCC and other malignancies acquire an abnormal enzymatic activity allowing them to convert α -ketoglutarate (α KG) to 2-hydroxyglutarate (2HG), which inhibits the activity of multiple α KG-dependent dioxygenases, and results in alterations in cell differentiation, survival, and extracellular matrix maturation^{6–10}. However, the molecular pathways by which IDH mutations lead to tumour formation remain unclear. Here we show that mutant IDH blocks liver progenitor cells from undergoing hepatocyte differentiation through the production of 2HG and suppression of HNF-4 α , a master regulator of hepatocyte identity and quiescence. Correspondingly, genetically engineered mouse models expressing mutant IDH in the adult liver show an aberrant response to hepatic injury, characterized by HNF-4 α silencing, impaired hepatocyte differentiation, and markedly elevated levels of cell proliferation. Moreover, *IDH* and *Kras* mutations, genetic alterations that co-exist in a subset of human IHCCs^{4,5}, cooperate to drive the expansion of liver progenitor cells, development of premalignant biliary lesions, and progression to metastatic IHCC. These studies provide a functional link between IDH mutations, hepatic cell fate, and IHCC pathogenesis, and present a novel genetically engineered mouse model of IDH-driven malignancy.

Gain-of-function *IDH1*/*IDH2* mutations occur in ~25% of IHCCs^{1,3–5}, a liver malignancy that exhibits bile duct differentiation, but have not been identified in hepatocellular carcinomas (HCCs), which exhibit hepatocyte differentiation (<http://www.sanger.ac.uk/cosmic>). To examine the role of IDH mutations in liver tumorigenesis we isolated mouse hepatoblasts, which are embryonic progenitors that give rise to hepatocytes and bile duct cells and correspond with adult liver progenitors^{11,12}. Hepatoblasts expressing mutant *IDH1* (R132C, R132H) or *IDH2* (R140Q, R172K) produced increased 2HG, but exhibited morphology and proliferation rates indistinguishable from vector and IDH wild-type controls (Extended Data Fig. 1a–d). However, unlike control hepatoblasts, which underwent hepatocyte differentiation when transferred from collagen-coated plates to uncoated plates¹³, forming hepatocyte clusters, decreasing proliferation, and activating a large program of hepatocyte-specific genes including *Adhl* and *Aldob*, IDH-mutant cells were refractory to differentiation (Fig. 1a–d and Extended Data Fig. 1e–g). *IDH1* (R132C) and *IDH2* (R172K) caused the most pronounced effects, correlating with relative 2HG levels. Treatment of R132C-expressing hepatoblasts with AGI-5027 (also known as ML309), a specific inhibitor of

mutant *IDH1* (ref. 14), attenuated 2HG production (Extended Data Fig. 1h) and restored hepatocyte differentiation (Fig. 1e, f). Conversely, differentiation of wild-type hepatoblasts was counteracted by (R)- or (S)-2HG octyl esters (Extended Data Fig. 1i, j). In contrast to the complete inhibition of hepatocyte differentiation, mutant IDH did not impair biliary differentiation of hepatoblasts in matrigel (Fig. 1g and Extended Data Fig. 1k, l). Thus, mutant IDH specifically blocks hepatocyte lineage progression through 2HG production.

To uncover the molecular program underlying these defects, we examined the effect of mutant *IDH1/2* on the transcriptome of hepatoblasts grown on collagen. Transcriptional profiles of mutant *IDH1* and *2* clustered together and gene set enrichment analysis (GSEA) demonstrated reduced expression of targets of HNF-4 α —a master transcriptional regulator of hepatocyte differentiation¹²—and of HNF-1 α , which acts downstream of HNF-4 α ¹⁵. Moreover, canonical HNF-4 α - and HNF-1 α -binding sites were strongly enriched at the promoters of differentially expressed genes (Fig. 2a and Extended Data Fig. 2a–c).

HNF-4 α has multiple isoforms expressed from separate promoters. The P2 promoter (encoding HNF-4 α (7–9)) is active in hepatoblasts and gradually extinguished in adult hepatocytes, whereas the P1 promoter (encoding HNF-4 α (1–6)) is hepatocyte-specific¹⁶. *Hnf4a*(7–9) messenger RNA and protein were reduced in IDH-mutant hepatoblasts, as was expression of HNF-4 α targets (Extended Data Fig. 2d–g). Moreover, under hepatocyte differentiation conditions, mutant IDH completely inhibited the pronounced induction of HNF-4 α (1–6) and its target—*OCLN*—that is observed in control cells (Fig. 2b and Extended Data Fig. 2h). Mutant IDH or octyl-2HG treatment blocked *Hnf4a*(1–6) mRNA induction, whereas AGI-5027 restored *Hnf4a*(1–6) levels in R132C-expressing cells (Fig. 2c and Extended Data Fig. 2i–k). Histone H3 lysine 4 trimethylation (H3K4me3) is associated with active transcription and was specifically reduced at the P1 promoter in R132C hepatoblasts, consistent with the observed silencing of *Hnf4a*(1–6), whereas the repressive marks, H3K27me3 and H3K9me3, were unaffected (Extended Data Fig. 2l, m and data not shown). Importantly, HNF-4 α knockdown impaired hepatocyte differentiation of wild-type hepatoblasts without inhibiting biliary differentiation, whereas ectopic HNF-4 α expression rescued differentiation of IDH-mutant cells (Fig. 2d–h and Extended Data Figs 2n, o, 3a, b). Thus, mutant IDH alters the epigenetic state of the P1 promoter—through targeting either direct regulators of the locus or more upstream factors—and prevents induction of *Hnf4a*(1–6), thereby blocking hepatocyte lineage progression. Notably, HNF-4 α and HNF-1 α have critical antiproliferative and tumour suppressor functions in the adult

¹Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts 02114, USA. ²Department of Medical Oncology, Dana-Farber Cancer Institute, Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. ³HCC Translational Research Laboratory, Barcelona-Clinic Liver Cancer Group, Liver Unit, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hospital Clínic, University of Barcelona, Catalonia 08036, Spain. ⁴Mount Sinai Liver Cancer Program, Division of Liver Diseases, Dept of Medicine, Icahn School of Medicine at Mount Sinai, New York 10029, USA. ⁵Gastrointestinal Surgery and Liver Transplantation Unit, National Cancer Institute, and Department of Experimental Oncology, Milan 20133, Italy. ⁶Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center, Kansas City, Kansas 66160, USA. ⁷University of Rochester Medical Center, Rochester, New York 14642, USA. ⁸Agios Pharmaceuticals, Cambridge, Massachusetts 02139, USA. ⁹Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia 08010, Spain. ¹⁰University of Barcelona, Catalonia 08036, Spain. ¹¹Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA.

*These authors contributed equally to this work.

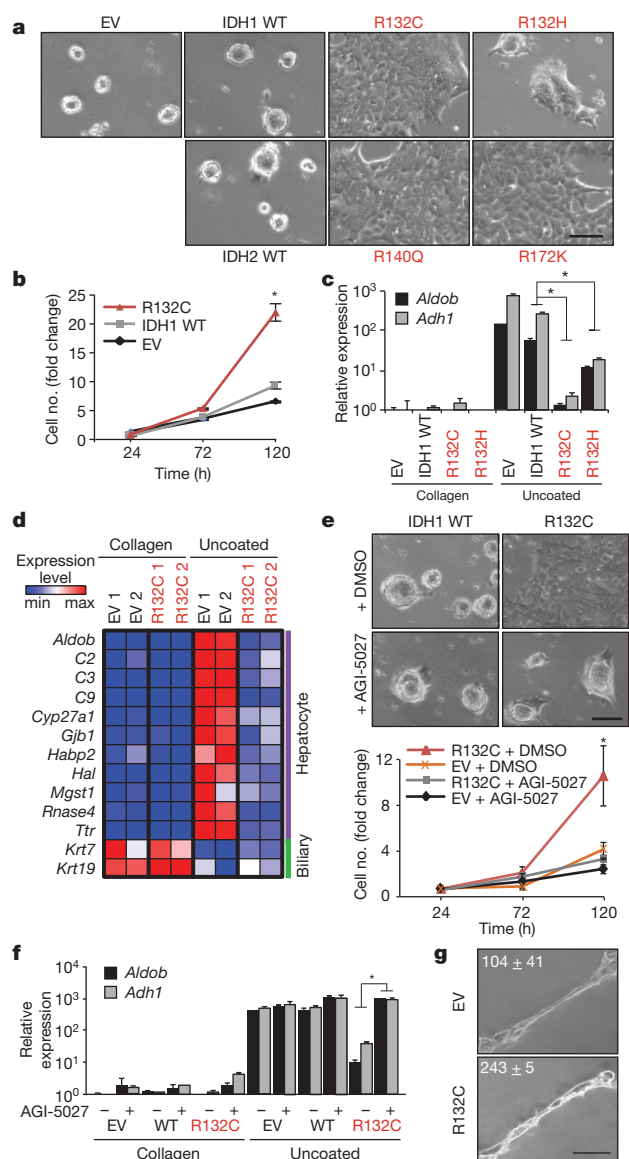


Figure 1 | Mutant IDH blocks hepatocyte differentiation. **a–d**, Hepatoblasts expressing empty vector (EV) or the indicated *IDH* alleles at physiological levels (see Extended Data Fig. 1a), were transferred to uncoated plates to induce hepatocyte differentiation. **a**, Hepatocyte sphere formation. WT, wild type. **b, c**, Proliferation (**b**) and hepatocyte marker expression (**c**) (quantitative polymerase chain reaction with reverse transcription (qRT-PCR)). **d**, Heat map of hepatocyte and biliary gene expression. **e, f**, Hepatoblasts treated with 2.5 μ M AGI-5027 or dimethylsulphoxide (DMSO) vehicle. **e**, Hepatocyte sphere formation (top), proliferation (bottom). **f**, Hepatocyte marker expression. **g**, Hepatoblasts in Matrigel assessed for biliary differentiation. Tubular structures per 6 cm dish \pm standard deviation (s.d.) are quantified. * $P < 0.05$. Scale bars, 100 μ m.

liver^{17–20}, suggesting the relevance of this pathway to mutant-IDH-mediated tumorigenesis and prompting us to extend our studies *in vivo*.

We generated transgenic mice with doxycycline-inducible expression of IDH2(R140Q) or IDH2(R172K) (Tet-R140Q, Tet-R172K strains) specifically in adult hepatocytes—R140Q was detected in virtually all hepatocytes and R172K showed more scattered expression, and liver 2HG levels were elevated (Extended Data Figs 4a–d, 5a). Since mutant IDH blocks liver progenitors from undergoing hepatocyte differentiation *in vitro*, we sought to address whether it acts analogously *in vivo* to specifically override differentiation from a progenitor cell state, or conversely, whether it broadly alters homeostasis of mature hepatocytes. Although normally quiescent, the liver has extensive regenerative

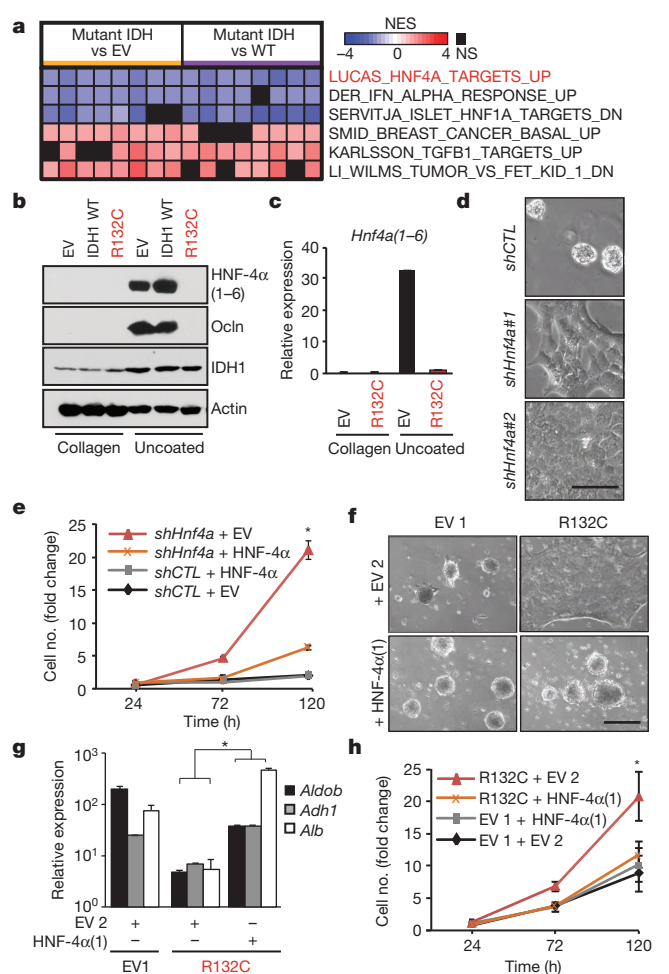


Figure 2 | Mutant IDH blocks hepatocyte differentiation by silencing HNF-4 α . **a**, Heat map of GSEA showing top-ranked gene sets distinguishing IDH1(R132C) or IDH2(R172K) from wild-type (WT) or empty vector (EV) control hepatoblasts (pairwise analysis; replicates for each condition; see Methods). NES, normalized enrichment score; NS, not significant. **b, c**, Hepatoblasts analysed by immunoblot (**b**) and qRT-PCR (**c**). **d, e**, Analysis of wild-type hepatoblasts expressing the indicated short hairpin (sh)RNAs (uncoated plates). CTL, control. **d**, Hepatocyte sphere formation. **e**, Proliferation of shRNA-expressing hepatoblast cells co-expressing EV or shRNA-resistant *Hnf4a(1)* complementary DNA. **f–h**, Control and R132C-expressing hepatoblasts co-expressing vector control (EV2) or HNF-4 α , grown on uncoated plates. **f**, Hepatocyte sphere formation. **g**, Hepatocyte gene expression. **h**, Proliferation. Scale bars, 100 μ m (**d**), 250 μ m (**f**). * $P < 0.05$.

capacity after injury involving replication of mature hepatocyte and biliary cells, or activation of bipotential progenitors (oval cells) that may arise from either lineage^{11,21}. In the absence of injury, Tet-R140Q mice were healthy for up to 48 weeks, and had normal liver histology, marker expression, proliferation, and liver function (Fig. 3d, Extended Data Fig. 5b and data not shown). By contrast, pronounced defects in restoration of hepatocyte differentiation were observed in mice fed a diet containing 3,5-diethoxycarbonyl-1,4-dihydrocollidin (DDC) for 5 days and then switched to a normal diet for 3 weeks (Fig. 3a), a protocol causing hepatocyte cell death and transient oval cell activation^{21,22}. Hepatocyte markers including HNF-4 α were downregulated 3–10-fold, while biliary markers were unchanged, and proliferation was increased >40-fold relative to wild-type controls (Fig. 3b–d). Despite this depletion of mature hepatocytes, no changes were seen in parameters of liver function (Extended Data Fig. 5c, d and data not shown), consistent with the persistence of hepatocytes surviving short-term DDC treatment and the established capacity of reduced hepatocyte numbers to maintain normal physiology.

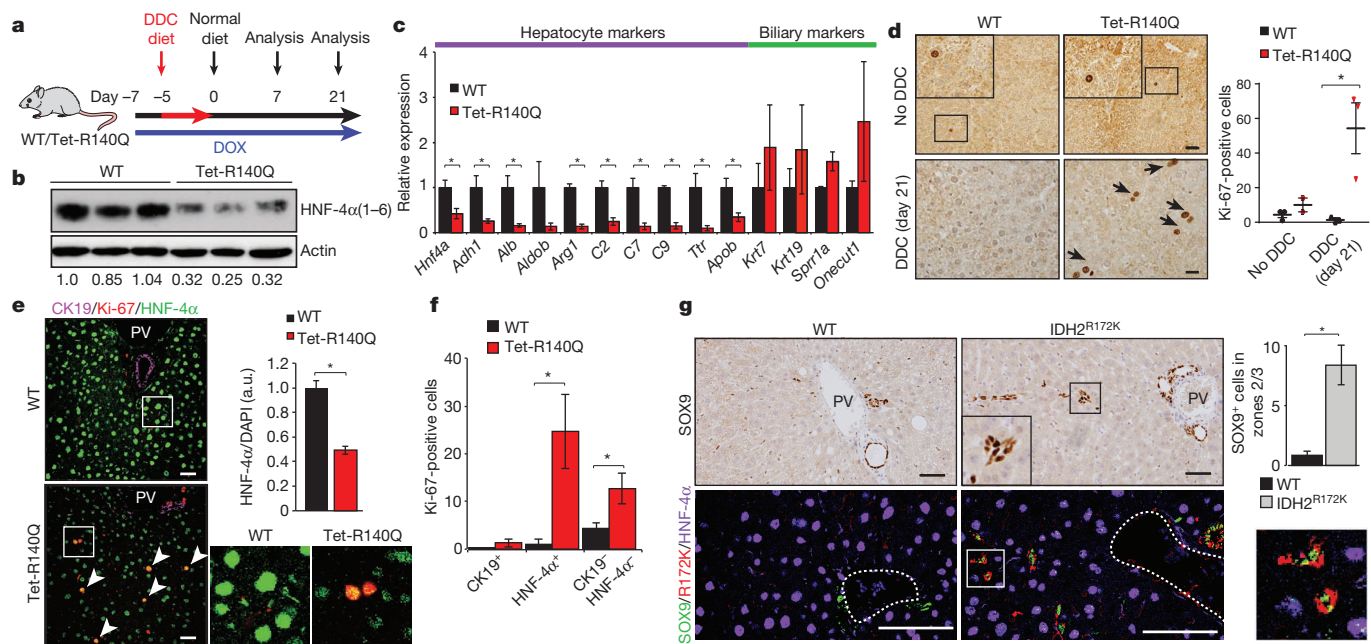


Figure 3 | Mutant IDH inhibits hepatocyte differentiation and quiescence of liver progenitors. **a**, Schematic of DDC study in Tet-R140Q (Tet-R140Q; Alb-Cre; Rosa26-LSL-rtTA) and wild-type (WT) littermate controls (Alb-Cre; Rosa26-LSL-rtTA). **b–f**, Livers at day 21. **b**, Immunoblot (HNF-4 α (1–6): actin is quantified). **c**, qRT-PCR. **d**, Ki-67 staining. Chart shows Ki-67⁺ cells per 20 high-powered fields. **e**, **f**, Immunofluorescence analysis. Graph depicts mean fluorescence intensity of HNF-4 α :4',6-diamidino-2-phenylindole (DAPI) (125 cells per group were scored). Inset: $\times 3$ magnification of boxed regions. a.u., arbitrary units. **f**, Quantification of Ki-67⁺ cells co-staining for the

indicated markers ($N = 3$ mice per group, 5 high-powered fields per mouse). **g**, Immunohistochemistry (top) and immunofluorescence (bottom) of wild-type (Alb-Cre) and IDH2^{R172K} (Alb-Cre; LSL-IDH2^{R172K}) livers at 20 months. Note accumulation of SOX9⁺ cells located $>25 \mu\text{m}$ away from bile duct or portal structures (dashed line), which express IDH2(R172K) and lack HNF-4 α . Inset: $\times 3$ magnification. Chart shows quantification. $N = 3$ mice per group; 4 high-powered images per mouse were scored. PV, portal vein. Error bars show mean \pm standard error of the mean (s.e.m.). Scale bars, 20 μm (**d**), 50 μm (**e**, **g**).

Serial analyses of wild-type and Tet-R140Q livers revealed comparable numbers of proliferating periductal HNF-4 α ⁺/CK19⁺ oval cells at 1 week, and resolution of this population after 3 weeks (Extended Data Fig. 6a, b and Fig. 3e, f). However, Tet-R140Q livers exhibited prominent induction of proliferating non-periductal cells with hepatocyte morphology but reduced or absent HNF-4 α expression (HNF-4 α ⁺ and HNF-4 α [−]/CK19⁺), which persisted after 3 weeks (Fig. 3e, f and Extended Data Fig. 6a–d). Similar but more tempered phenotypes were seen in Tet-R172K mice, consistent with focal transgene expression (Extended Data Fig. 6e, f). Thus, mutant IDH specifically blocks restoration of hepatocyte differentiation after acute liver injury, leading to aberrant proliferation in the hepatic parenchyma. As HNF-4 α levels increase in liver progenitors undergoing hepatocyte commitment^{23,24}, these proliferating HNF-4 α -low/absent cells appear to be committed progenitors (derived from dedifferentiated hepatocytes or oval cells) whose differentiation to hepatocytes is subverted by mutant IDH2. The transgene was not expressed in SOX9⁺ oval cells, precluding assessment of effects on these earlier progenitors.

We developed an additional transgenic strain (IDH2^{R172K}) expressing human IDH2(R172K) in SOX9⁺ biliary cells but not in hepatocytes (Fig. 3g and Extended Data Fig. 7a, b). These animals displayed normal liver histology, gene expression, and proliferation at 3 months (Extended Data Fig. 7c and data not shown). However, by 20 months of age there was pronounced accumulation of HNF-4 α [−]/SOX9⁺ oval cells expressing IDH2(R172K) $>25 \mu\text{m}$ away from any bile duct or portal structure (zones 2/3) (Fig. 3g). Collectively, our results suggest that mutant IDH2 abrogates differentiation of adult progenitors (activated spontaneously during ageing or by injury), specifically blocking hepatocyte lineage progression.

These findings are concordant with our observations in hepatoblasts and suggest that failed HNF-4 α induction could contribute to the liver phenotypes caused by mutant IDH *in vivo*. In this regard, acute *Hnf4a* deletion in hepatocytes is reported to provoke hepatocyte differentiation

defects and, upon diethylnitrosamine (DEN) treatment, oval cell accumulation and formation of tumours showing HCC and IHCC morphology²⁰. To expand upon potential parallels with IDH-mutant mice, we characterized these phenotypes in further depth. Importantly, the HCC lesions in DEN-treated *Hnf4a* conditional knockout mice were uniformly HNF-4 α ⁺, indicating that they arose from cells that escaped deletion of the locus (Extended Data Fig. 8a). By contrast, SOX9⁺ oval cells and CK19⁺ IHCCs were HNF-4 α [−] and were never observed in DEN-treated controls (Extended Data Fig. 8b, c). Thus, HNF-4 α ablation in DEN-treated livers drives progenitor expansion and progression specifically to IHCC. These findings establish *Hnf4a* as an IHCC tumour suppressor and are consistent with HNF-4 α acting downstream of mutant IDH in liver growth control.

The capacity of mutant IDH to silence *Hnf4a* and impair differentiation would be expected to confer sensitivity to transformation by additional oncogenic lesions. In this regard, IDH and KRAS mutations exist concurrently in human IHCC^{4,5} and acute myeloid leukaemia²⁵. Previously, we showed that Kras(G12D) expression in mouse liver causes mixed IHCC/HCC with long latency, a phenotype accelerated by TP53 deletion²⁶. Intercrossing LSL-IDH2^{R172K}, LSL-Kras^{G12D} and Alb-Cre mice revealed dramatic oncogenic cooperation, with six out of six Alb-Cre; LSL-IDH2^{R172K}; LSL-Kras^{G12D} (IDH2^{R172K}; Kras^{G12D}) animals developing poor body condition and palpable liver tumours between 33 and 58 weeks (mean 47.3 weeks; Fig. 4a). Multifocal liver masses with splenic invasion and peritoneal metastases were observed and demonstrated to be IHCC by histopathological analysis, CK19 staining, and lack of reactivity for the hepatocyte/HCC marker Hep Par1 (Fig. 4b). By contrast, only one out of seven Alb-Cre; LSL-Kras^{G12D} (Kras^{G12D}) mice sustained a tumour by 70 weeks (mean survival = 81.6 weeks), and solely HCCs were detected (Extended Data Fig. 8d). Tumour 2HG levels in IDH2^{R172K}; Kras^{G12D} mice were comparable to those in IDH-mutant human IHCC¹ (Extended Data Fig. 8e).

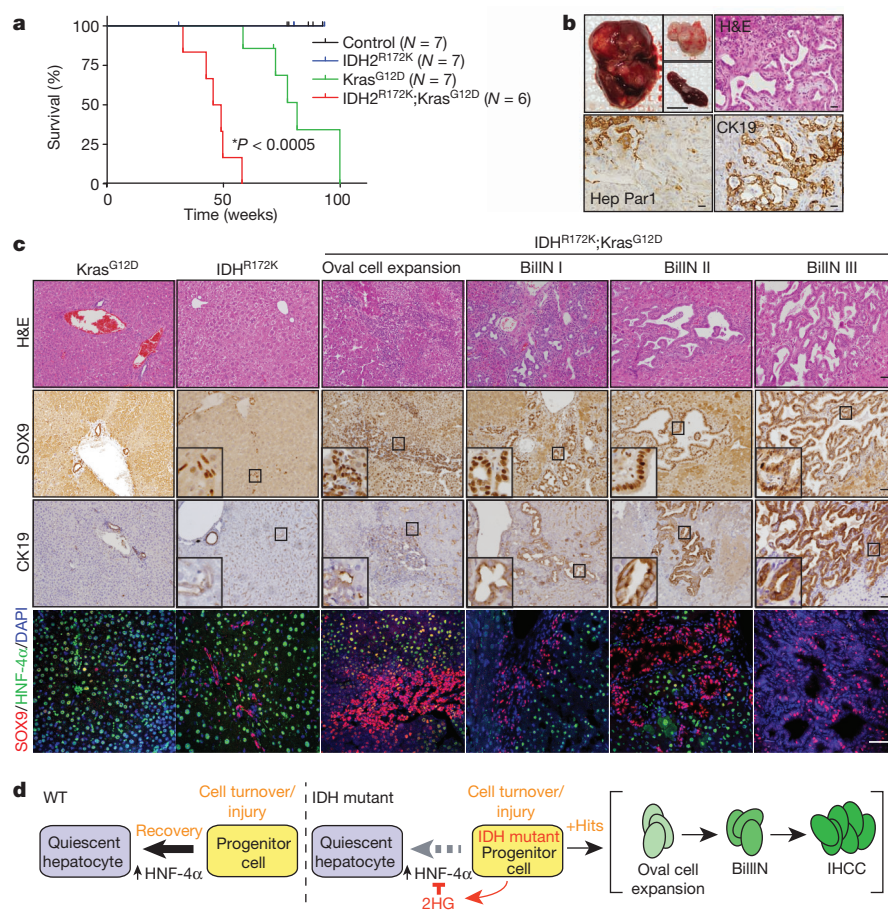


Figure 4 | Mutant IDH cooperates with Kras(G12D) to drive liver progenitor cell expansion and multi-step IHCC pathogenesis. **a**, Kaplan–Meier analysis showing time until signs of illness necessitated euthanasia. All animals euthanized had liver tumours. **b**, Top left, representative IDH2^{R172K};Kras^{G12D} (Alb-Cre;LSL-IDH2^{R172K};LSL-Kras^{G12D}) tumour, and peritoneal and spleen metastases (insets). Top right, haematoxylin and eosin (H&E) staining showing IHCC histology. Bottom, the tumour is Hep Par1[−] and CK19⁺ while adjacent hepatocytes stain Hep Par1⁺ and CK19[−]. **c**, The livers of IDH2^{R172K};Kras^{G12D} animals exhibit oval cell expansion and increasing grades of biliary intraepithelial neoplasias (BilIN), which stain SOX9⁺. CK19 levels increase with higher grade lesions. Immunofluorescence analysis reveals focal accumulation of SOX9⁺ oval cells in IDH2^{R172K} livers and pronounced oval cell expansion in IDH2^{R172K};Kras^{G12D} livers. **d**, Model for mutant IDH in IHCC pathogenesis. Scale bars, 1 cm (**b**, top left), 50 μ m (**b**, **c**).

In humans, IHCC is thought to develop from precursor lesions, including biliary intraepithelial neoplasias². Aberrant oval cell proliferation also precedes liver cancer development in murine models²⁷. Notably, in all IDH2^{R172K};Kras^{G12D} mice analysed ($N = 6$), the adjacent liver exhibited oval cell expansion and a graded series of biliary intraepithelial neoplasia-like lesions, which were SOX9⁺/CK19⁺, and expressed the IDH2^{R172K} transgene at physiological levels (Fig. 4c and Extended Data Figs 9a, b, 10a). Importantly, neither Kras^{G12D} nor Kras^{G12D};p53^{Lox/+} livers²⁶ showed oval cell expansion, and biliary intraepithelial neoplasias were found in only two of eight Kras^{G12D};p53^{Lox/+} mice. Thus, IDH2 (R172K) and Kras(G12D) cooperatively incite activation of hepatic progenitors and multistage IHCC progression.

We present a novel genetically engineered mouse model driven by mutant IDH that exhibits the hallmarks of human IHCC. Mutant IDH blocks hepatocyte differentiation from progenitors *in vitro*, and regulates hepatic regeneration, progenitor cell expansion, and IHCC pathogenesis *in vivo*. Collectively, our findings are consistent with a model whereby mutant IDH subverts the HNF-4 α -mediated hepatocyte differentiation/quiescence program in proliferating hepatocytes or bipotential progenitors, creating a persistent pre-neoplastic state primed for transformation by additional oncogenic mutations, and leading to adenocarcinoma specifically with a biliary phenotype (Fig. 4d). While lineage-tracing studies are required to fully define the impact of IDH mutations on different liver cell types, this model predicts that IDH-mutant IHCC may have progenitor-like features, and that IDH mutations occur early in disease pathogenesis as observed in glioblastoma and acute myeloid leukaemia^{28,29}. Accordingly, GSEA of human IHCCs revealed that the IDH-mutant subset (17/107 cases) have strong enrichment of a hepatic stem cell signature (Extended Data Fig. 10b).

Defining the biological functions of mutant IDH is of immediate importance as IHCCs are resistant to current treatment. Our genetically engineered mouse model provides an autochthonous context for

'co-clinical trials' to help determine the best way to deploy IDH inhibitors and identify biomarkers of response, thereby informing clinical decision-making in this deadly disease.

METHODS SUMMARY

Cell culture. Hepatoblasts were prepared from wild-type mice at embryonic day 14 and subjected to hepatocyte and bile duct differentiation as described¹³. For proliferation assays, cells were plated in duplicate on collagen-coated or uncoated 6-well plates. Cells were trypsinized and counted by trypan-blue exclusion.

Mice. Mice were housed in pathogen-free animal facilities. Studies were approved by the Subcommittee on Research Animal Care at Massachusetts General Hospital (protocol 2005N000148). The Tet-R140Q, Tet-R172K and IDH2^{R172K} transgenic mice and other strains are described in Methods. For DDC experiments, mice were maintained on 200 μ g ml^{−1} doxycycline-containing water. Mice were fed with 0.1% DDC-containing chow (F4643, Bio-serv) from days 2–7. Mice were then switched to normal chow until harvest. DEN experiments were performed as described²⁰.

Statistics. Results are expressed as mean \pm s.d. unless otherwise specified. Significance was analysed using two-tailed Student's *t*-test. A *P* value of less than 0.05 was considered statistically significant.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 September 2013; accepted 2 May 2014.

Published online 2 July 2014.

1. Borger, D. R. *et al.* Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist* **17**, 72–79 (2012).
2. Hezel, A. F., Deshpande, V. & Zhu, A. X. Genetics of biliary tract cancers and emerging targeted therapies. *J. Clin. Oncol.* **28**, 3531–3540 (2010).
3. Razumilava, N. & Gores, G. J. Cholangiocarcinoma. *Lancet* [http://dx.doi.org/10.1016/S0140-6736\(13\)61903-0](http://dx.doi.org/10.1016/S0140-6736(13)61903-0) (2014).
4. Voss, J. S. *et al.* Molecular profiling of cholangiocarcinoma shows potential for targeted therapy treatment decisions. *Hum. Pathol.* **44**, 1216–1222 (2013).

5. Wang, P. *et al.* Mutations in isocitrate dehydrogenase 1 and 2 occur frequently in intrahepatic cholangiocarcinomas and share hypermethylation targets with glioblastomas. *Oncogene* **32**, 3091–3100 (2013).
6. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744 (2009).
7. Losman, J. A. & Kaelin, W. G. Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes Dev.* **27**, 836–852 (2013).
8. Lu, C. *et al.* IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* **483**, 474–478 (2012).
9. Sasaki, M. *et al.* IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. *Nature* **488**, 656–659 (2012).
10. Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α -ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30 (2011).
11. Michalopoulos, G. K. Principles of liver regeneration and growth homeostasis. *Compr Physiol* **3**, 485–513 (2013).
12. Si-Tayeb, K., Lemaigre, F. P. & Duncan, S. A. Organogenesis and development of the liver. *Dev. Cell* **18**, 175–189 (2010).
13. Strick-Marchand, H. & Weiss, M. C. Inducible differentiation and morphogenesis of bipotential liver cell lines from wild-type mouse embryos. *Hepatology* **36**, 794–804 (2002).
14. Davis, M. I. *et al.* Biochemical, cellular and biophysical characterization of a potent inhibitor of mutant isocitrate dehydrogenase IDH1. *J. Biol. Chem.* **289**, 13717–13725 (2014).
15. Kuo, C. J. *et al.* A transcriptional hierarchy involved in mammalian cell-type specification. *Nature* **355**, 457–461 (1992).
16. Torres-Padilla, M. E., Fougere-Deschatrette, C. & Weiss, M. C. Expression of HNF4 α isoforms in mouse liver development is regulated by sequential promoter usage and constitutive 3' end splicing. *Mech. Dev.* **109**, 183–193 (2001).
17. Bluteau, O. *et al.* Bi-allelic inactivation of *TCF1* in hepatic adenomas. *Nature Genet.* **32**, 312–315 (2002).
18. Bonzo, J. A., Ferry, C. H., Matsubara, T., Kim, J. H. & Gonzalez, F. J. Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4 α in adult mice. *J. Biol. Chem.* **287**, 7345–7356 (2012).
19. Servitja, J. M. *et al.* Hnf1 α (MODY3) controls tissue-specific transcriptional programs and exerts opposed effects on cell growth in pancreatic islets and liver. *Mol. Cell. Biol.* **29**, 2945–2959 (2009).
20. Walesky, C. *et al.* Hepatocyte nuclear factor 4 α deletion promotes diethylnitrosamine-induced hepatocellular carcinoma in rodents. *Hepatology* **57**, 2480–2490 (2013).
21. Yanger, K. *et al.* Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes Dev.* **27**, 719–724 (2013).
22. Preisegger, K. H. *et al.* Atypical ductular proliferation and its inhibition by transforming growth factor beta1 in the 3,5-diethoxycarbonyl-1,4-dihydrocollidine mouse model for chronic alcoholic liver disease. *Lab. Invest.* **79**, 103–109 (1999).
23. Malato, Y. *et al.* Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration. *J. Clin. Invest.* **121**, 4850–4860 (2011).
24. Shin, S. *et al.* Foxl1-Cre-marked adult hepatic progenitors have clonogenic and bilineage differentiation potential. *Genes Dev.* **25**, 1185–1192 (2011).
25. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
26. O'Dell, M. R. *et al.* Kras(G12D) and p53 mutation cause primary intrahepatic cholangiocarcinoma. *Cancer Res.* **72**, 1557–1567 (2012).
27. Avruch, J., Zhou, D., Fitamant, J. & Bardeesy, N. Mst1/2 signalling to Yap: gatekeeper for liver size and tumour development. *Br. J. Cancer* **104**, 24–32 (2011).
28. Tefferi, A. *et al.* IDH1 and IDH2 mutation studies in 1473 patients with chronic, fibrotic- or blast-phase essential thrombocythemia, polycythemia vera or myelofibrosis. *Leukemia* **24**, 1302–1309 (2010).
29. Watanabe, T., Nobusawa, S., Kleihues, P. & Ohgaki, H. IDH1 mutations are early events in the development of astrocytomas and oligodendrogliomas. *Am. J. Pathol.* **174**, 1149–1153 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Mostoslavsky, L. Ellisen, A. Kimmelman, and members of the Bardeesy laboratory for valuable input. We also thank S. Thorgeirsson and J. Andersen for sharing unpublished data sets. This work was supported by grants from TargetCancer Foundation and the National Institutes of Health (R01CA136567-02 and P50CA1270003) to N.B. N.B. holds the Gallagher Endowed Chair in Gastrointestinal Cancer Research at Massachusetts General Hospital. S.K.S. is the recipient of a Cholangiocarcinoma Foundation/Conquer Cancer Foundation of ASCO Young Investigator Award, and an American Cancer Society Postdoctoral Fellowship (PF-13-294-01-TBG). C.A.P. is the recipient of a Canadian Institutes of Health Research postdoctoral fellowship. N.B., J.M.L. and D.S. are members of the Samuel Waxman Cancer Research Foundation Institute Without Walls. J.M.L. and D.S. are supported by the Asociación Española para el Estudio del Cáncer.

Author Contributions S.K.S. and C.A.P. contributed equally to the study. S.K.S., C.A.P., K.S.G., M.S.N. and S.G. carried out the experiments involving hepatoblasts and mouse models. K.N.R. and S.R. performed computational analysis on gene expression data. J.F. performed immunohistochemistry on tissue sections. E.A.A. and K.-K.W. assisted with the generation of the IDH mutant mice. V.D. analysed the histology from the murine liver specimens. C.W. and U.A. carried out the experiments involving the HNF-4 α knockout mice. D.S., H.C., O.M. and J.M.L. performed GSEA analysis on human IHCC samples. A.X.Z., A.F.H. and C.R.F. were involved in the study design. K.E.Y., K.S.S., J.T., J.P.-M. and C.G. developed and provided the AGI-5027 compound and measured 2HG in our samples. S.K.S., C.A.P. and N.B. designed the experiments and wrote the paper. N.B. supervised the studies. All authors discussed the results and commented on the manuscript.

Author Information Affymetrix Mouse 420Av2 DNA microarray data have been deposited in the Gene Expression Omnibus under accession number GSE57002. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.B. (Bardeesy.Nabeel@mgh.harvard.edu).

Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells

Zohar Shipony^{1*}, Zohar Mukamel^{1*}, Netta Mendelson Cohen¹, Gilad Landan¹, Elad Chomsky^{1,2}, Shlomit Reich Zeliger³, Yael Chagit Fried⁴, Elena Ainbinder⁴, Nir Friedman³ & Amos Tanay¹

Stable maintenance of gene regulatory programs is essential for normal function in multicellular organisms. Epigenetic mechanisms, and DNA methylation in particular, are hypothesized to facilitate such maintenance by creating cellular memory^{1–4} that can be written during embryonic development^{5,6} and then guide cell-type-specific gene expression⁷. Here we develop new methods for quantitative inference of DNA methylation turnover rates, and show that human embryonic stem cells preserve their epigenetic state by balancing antagonistic processes that add and remove methylation marks rather than by copying epigenetic information from mother to daughter cells. In contrast, somatic cells transmit considerable epigenetic information to progenies. Paradoxically, the persistence of the somatic epigenome makes it more vulnerable to noise, since random epimutations can accumulate to massively perturb the epigenomic ground state. The rate of epigenetic perturbation depends on the genomic context, and, in particular, DNA methylation loss is coupled to late DNA replication dynamics. Epigenetic perturbation is not observed in the pluripotent state, because the rapid turnover-based equilibrium continuously reinforces the canonical state. This dynamic epigenetic equilibrium also explains how the epigenome can be reprogrammed quickly⁸ and to near perfection⁹ after induced pluripotency.

According to the epigenetic persistence model, information on cellular state or cellular fate is stored on the chromosomal template and copied directly from mother to daughter cells during DNA replication. The paradigm for this model is the replication of symmetrically methylated CpG dinucleotides by the mammalian housekeeping methyltransferase DNMT1 (refs 10–12). According to an alternative model, epigenetic state is maintained through a dynamic balance between *trans*-acting factors and *cis*-epigenetic marks, and the maintenance of the state through replication is facilitated by the cytoplasmatic and nuclear concentration of the relevant *trans*-factors. Recently, the role of DNA methylation in forming epigenetic memory through persistent or dynamic mechanisms has been re-evaluated, given the characterization of pathways leading to active demethylation through hydroxylation (most notably in embryos or embryonic stem (ES) cells^{13,14}) and given observations suggesting the methylation state within cell populations is very heterogeneous and possibly noisy¹⁵.

To distinguish these memory models and quantify their relative contribution, we combined unique molecular identifiers (UMIs) with reduced representation bisulphite sequencing^{16,17} (UMI-bis; Extended Data Fig. 1a–i), thereby facilitating truly quantitative profiling of DNA methylation pattern distributions within cell populations. UMI-bis of ES cells and somatic fibroblasts populations reconfirmed the well-documented bipolar (on–off) average methylation landscape in these cellular states (Extended Data Fig. 2a–c). The intrapopulation methylation distributions reconstructed were, however, distinct. ES cells show more rare methylation patterns (which we define as noise) in loci that are dominated by methylated patterns. Fibroblasts, on the other hand, show fewer rare patterns, but more

cases of loci that are dominated by a specific partially methylated pattern (Extended Data Fig. 2d–h). These observations suggested that the DNA methylation dynamics, and more generally the mechanisms for transmitting epigenetic memory, are fundamentally different between ES cells and fibroblasts.

To quantify DNA methylation dynamics in ES cells we implemented an (epi)mutation accumulation approach¹⁸. We generated clonal cell populations ($n = 5 \times 10^4$ – 2×10^6) and assayed their methylation distributions while considering their defined single (for H1 male X chromosome) or double (for all other chromosomes) epiallelic origin. We reasoned that given the heterogeneity of the polyclonal ES-cell methylome, it would be easy to detect any trace of epigenetic persistence in this system by observing rare patterns in the population that become common (that is, approaching 50% of the data) in one of the clones. Despite this expectation, the frequencies of the most common pattern in each of our clonal populations (Fig. 1a and Extended Data Fig. 2i) showed perfect correspondence with their frequencies in the matched polyclonal population. We next clustered loci according to the frequencies of their unmethylated, methylated and partially methylated patterns, as well as according to their total frequency of rare patterns (noise) (Extended Data Fig. 3). We grouped clusters into nine classes according to average methylation. As shown in Fig. 1b, classes I–III comprise unmethylated loci and consist of homogeneous clusters that differ only in the frequency of rare patterns. These clusters reflect no notable distinction between clonal and non-clonal populations but do indicate that the level of methylation noise is a conserved property that can be reconstituted quickly after single-cell cloning. Loci in classes I and II show low noise and are typically found at promoters of active transcription start sites, while class III loci show high noise and are correlated with H3K27me3 occupancy (and, to a lesser extent, enhancer regions) (Fig. 1c).

In marked contrast to ES cells, the most frequent methylation patterns observed in clonal fibroblast populations often show up as rare patterns in the matching polyclonal population (Fig. 1d and Extended Data Fig. 2j), suggesting that fibroblasts retain a considerable amount of clonal epigenetic memory. Clustering of genomic loci according to the distribution of methylation patterns in six clones and the matched polyclonal population supports this idea (Fig. 1e) by identifying six clusters (class III) that are dominated by two epialleles in one clone (one unmethylated, the other partially methylated), but are symmetrically unmethylated in the population and all other clones. Also identified are groups of sites with low and medium levels of noise (classes I and II), as observed in ES cells, but not a cluster with high noise level. Furthermore, annotation of the observed low methylation clusters indicates that class III fibroblast loci are enriched with H3K27me3 marks in ES cells (Fig. 1f). As shown earlier, these regions are precisely those defined by high noise levels in ES cells. In conclusion, low methylation loci in ES cells are maintained without direct evidence for clonal memory, showing variable levels of noise that are correlated with their epigenetic context. Fibroblasts do not

¹Department of Computer Science and Applied Mathematics, and Department of Biological Regulation, Weizmann Institute of Science, Rehovot 76100, Israel. ²Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel. ³Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel. ⁴Department of Biological Services, Weizmann Institute of Science, Rehovot 76100, Israel.

*These authors contributed equally to this work.

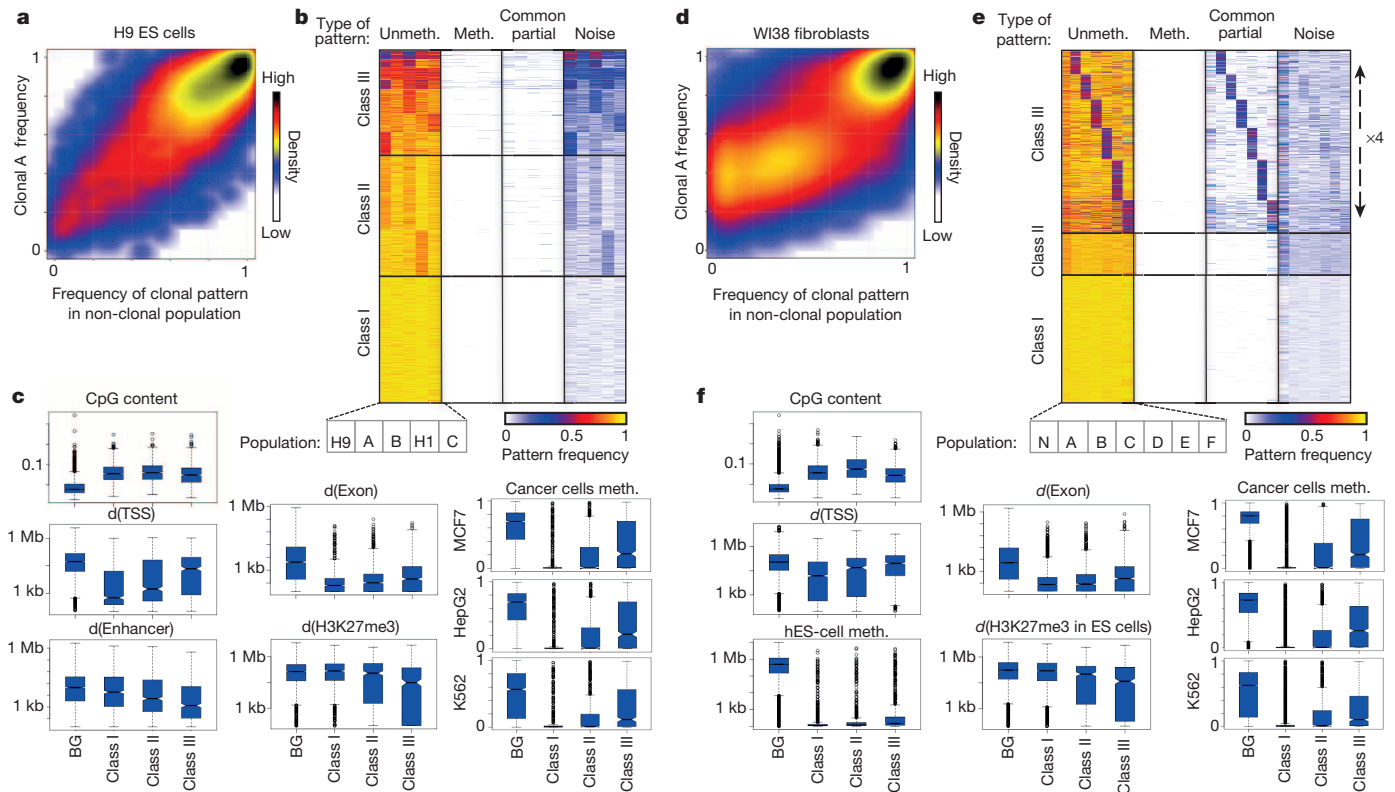


Figure 1 | Clonal memory and rapid turnover at low-methylation loci.

a, For each locus, the UMI-bis-derived frequency of the most abundant methylation pattern in one clonal H9 ES-cell population (y-axis) is compared to the frequency of the same pattern in the matching polyclonal population (x-axis). Clonal memory should theoretically be identified as off-diagonal cases in which high clonal frequency is observed for patterns with low polyclonal frequencies. **b**, We define the methylation distributions of each locus in H9 and H1 polyclonal populations and three clonal populations (A, B and C) using the frequencies of the unmethylated (Unmeth.), methylated (Meth.) and partially methylated common (Common partial) patterns (if they exist).

show high methylation noise at H3K27me3-linked regions, and sporadic, initially rare methylation in these regions is clonally transmitted. Interestingly, even though this noise propagation process is slow and does not change average methylation at the affected loci immediately, it is associated with a remarkable predisposition for developing hypermethylation in cancer cells and tumours^{19,20} (Fig. 1f).

Clusters of loci with high average methylation in clonal and non-clonal ES-cell populations showed that clonal epigenetic memory is not detectable in these regions (Fig. 2a), as observed for loci with low average methylation. In contrast, analysis of fibroblast data identified six groups of loci in which one clone replaced up to 50% of the methylated patterns with a partially methylated pattern (Fig. 2b, class VI), apparently memorizing a randomly selected partially methylated epiallele in the founding single cell. Also identified are groups of sites (Fig. 2b, class IV) in which one or more clones regain high frequency of the methylated pattern in place of a partially methylated pattern that is observed at around 50% in all other clonal and polyclonal populations. These dynamics suggest that loci with intermediate methylation in fibroblasts emerged in many cases through multiple independent and clonally transmitted epimutations that erode an initial perfectly methylated state. Indeed, loci with average methylation of 30–70% in fibroblasts (but not ES cells) show much greater variability between clones than within them (Fig. 2c), as measured by intra- and interclonal epipolymorphism. This observation suggests that intermediate methylation loci are dominated by a different pattern in each single cell and that the progenies of each such cell maintain this pattern while adding some additional noise to it. Clusters of

The total frequency of rare patterns defines the locus epigenetic noise. Shown here are clusters with low average methylation and variable (yet consistent) noise levels (classes I–III). **c**, Boxplots show the genomic enrichments of loci within the methylation pattern classes. Also shown are average methylation levels in three cancer cell lines. **d**, distance from genomic feature. BG, background **d**–f, Analogous to **a**–c, but using data from six clonal somatic fibroblasts populations (A–F). h, human; N, non-clonal. For boxplots, boxes span the data interquartile range (ICR, between the 1st and 3rd quartile); whiskers span the data range but extend no more than 1.5 of the ICR above and below the box; outliers (if present) are omitted.

intermediate methylation loci in fibroblasts (Fig. 2e, classes VII and VIII) show that methylation in a large fraction of these sites is distributed over a large number of patterns that are observed at a high frequency in at least one clone (numbered 1–6 in Fig. 2e). Loci with intermediate methylation in ES cells (Fig. 2d) are generally rare, and do not reflect such behaviour. In conclusion, the methylation dynamics at intermediate methylation regions in fibroblasts demonstrate how clonal persistence and epimutations can together promote the emergence of epigenetic heterogeneity.

Epigenomic annotation of high- and intermediate-methylation clusters suggested that persistent loci in methylated regions, and partially methylated regions in general, are enriched in large epigenomic domains that replicate late in S phase (Fig. 2f) or are lamina associated^{21,22} (Extended Data Fig. 4). Interestingly, the persistent dynamics of DNA methylation, as detected in the clonal analysis of the fibroblast system, are correlated with pronounced variability in the DNA methylation content in different cancer cell lines (Fig. 2g). To further explore this variability, we generated single-cell clonal populations from human HepG2, MCF7 and K562 cells. In all of these systems we observed notable clusters showing persistent clonal methylation. The data suggest that in MCF7 cells, a shift towards a regime favouring methylation gain over loss gives rise to near complete methylation of regions that are heterogeneous in fibroblasts. The opposite regime gives rise to global loss of methylation at the same regions in the K562 cell line. HepG2 cells maintain a similar balance to fibroblasts (Fig. 2g and Extended Data Figs 5, 6). These observations support a model in which methylation state is determined by the balance between stochastic processes of gaining and losing methyl marks, and this balance can change

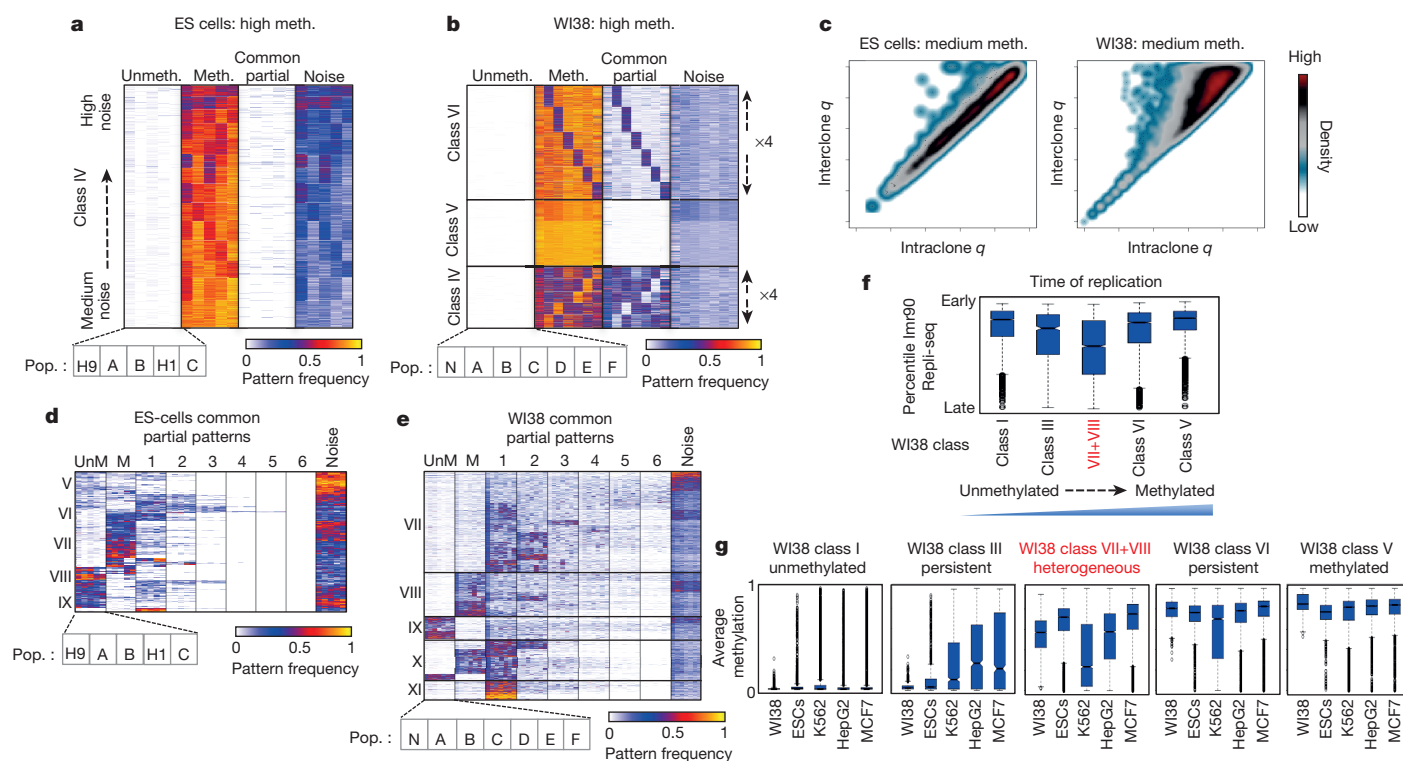


Figure 2 | Clonal memory and the emergence of a heterogeneous methylation distributions. **a**, ES-cell pattern clusters of loci with high methylation average, displayed as in Fig. 1b. **b**, Fibroblasts pattern clusters of loci with high average methylation. **c**, Density map comparing intraclonal and interclonal epipolymorphism (q) for fragments with sufficient coverage in ES cells (left) and fibroblasts (right). **d**, Pattern clusters for regions with intermediate average methylation in ES cells. Patterns 1–6 represent, for each locus, all common partial patterns that could be found in the clonal populations. M, methylated; UnM, unmethylated. **e**, Similar to **d**, showing

fibroblast data. **f**, The boxplots show the distributions of replication time for loci classified according to the methylation distributions shown above. **g**, The boxplots show the distributions of average methylation in ES cells (ESCs), fibroblasts and cancer cell lines for loci classified according to the fibroblast pattern classes. For boxplots, boxes span the data interquartile range (ICR, between the 1st and 3rd quartile); whiskers span the data range but extend no more than 1.5 of the ICR above and below the box; outliers (if present) are omitted.

in a cell-type specific fashion. Analysis of primary mouse CD8⁺ T cells (Extended Data Fig. 7a–e) shows that, much like in human fibroblasts, persistent unmethylated loci in mouse T cells are enriched with polycomb occupancy in mouse ES cells, and that a large group of loci develop heterogeneous methylation distributions at late replicating, repressive domains. A similar observation is derived from analysis of methylation loss in adult human tissues compared to ES cells (Extended Data Fig. 7f, g).

Relying on the relative stability of the methylation landscapes in non-ES-cell populations, we next developed an algorithm for probabilistic inference of the ancestral methylation states of the two epialleles founding the population at each locus, and the series of methylation gain and loss events that transform these epialleles into those observed in the clonal population (Extended Data Fig. 8). This approach enabled estimation of methylation gain and loss rates across the genome. By pooling data from the fibroblast clones we inferred gain and loss rates around active gene promoters (Fig. 3a) and enhancers (Fig. 3b)^{23,24}, supporting the existence of a dominant process of methylation loss in these hypomethylated regions. A more complex behaviour was observed on well-organized nucleosome arrays that are anchored by CTCF binding (Fig. 3c)²⁵, where fluctuating rates of methylation gain and loss rates are correlated with nucleosome occupancy and fluctuating average methylation levels. Importantly, gain and loss rates are not a simple function of average methylation in these regions, and higher or lower overall turnover rates can be observed in regions with similar average methylation. This suggests that the balance between these rates, and not their absolute intensity, is what defines methylation average in the long run.

We next performed global stratification analysis to understand the relationships between methylation turnover and replication time, lamina

association, and nucleosome occupancy^{23,24,26} (Fig. 3d). This analysis showed that methylation loss rates are remarkably dependent on replication time, but not on lamina or nucleosome occupancy. In contrast, methylation gain rates are positively correlated with nucleosome occupancy, and are also (independently) stronger in lamina-associated sites. Inference of turnover rates from additional fibroblast clonal populations sampled at three consecutive time points confirmed the robustness of the turnover estimates (Extended Data Fig. 9a–e). Turnover rate analysis of K562 cells (Extended Data Fig. 9d) demonstrated much weaker lamina- and nucleosome-dependent gain rates in these cells, but also indicated very similar trends associated with methylation loss rates.

Finally, we screened the H9 ES-cell (XX karyotype) methylome for pattern distributions showing high methylation variance, aiming to explore the methylation dynamics in regions that are not dominated by complete ES-cell methylation or a complete lack of it. This resulted in 417 regions on the X chromosome and 281 regions on autosomes. Such high pattern variance can be explained by conserved allele-specific methylation (differential methylation of chromosomes within the same cell), a subpopulation structure (differential methylation of both chromosomes in a subpopulation), or by more specific dynamics that generate fluctuations between an ‘open’ and ‘closed’ chromatin state at particular chromosomal regions, driving rapid demethylation and remethylation, respectively (Extended Data Fig. 10a–c). Further analysis of high-variance X-linked loci showed that these are highly conserved in clonal populations (Fig. 4a), indicating likely monoallelic methylation on the inactive X that is conserved at the single-cell level and therefore maintained in clones. On the other hand, autosomal high-variance loci were less conserved clonally (Fig. 4a), suggesting that in addition to classical monoallelic

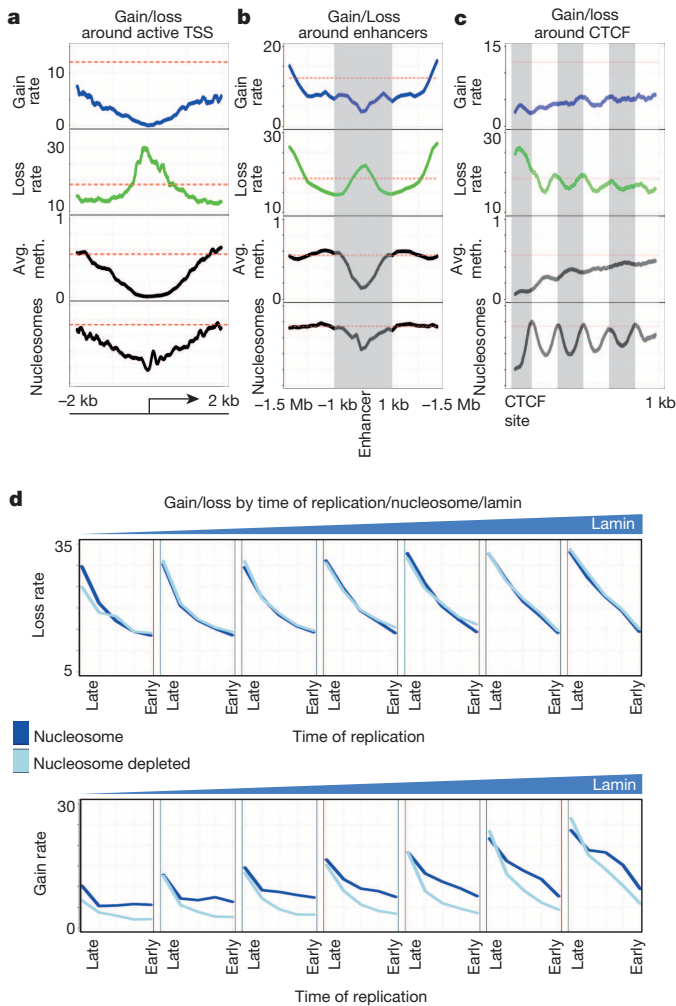


Figure 3 | Turnover rates are correlated with local epigenomic context. **a–c**, Shown are rates of methylation gain and loss that were estimated by pooling data from six clonal populations and stratifying on the basis of genomic context. Also shown are average methylation (Avg. meth.) and average nucleosome occupancy on the same sites. **d**, Shown are methylation gain and loss rates stratified according to nucleosome occupancy (light blue, depleted; blue, occupied), time of replication (x-axis of each panel), and nuclear lamina interaction (different panels organized from low (left) to high (right)).

methylation, some H9 loci may be explicable by the existence of differentially methylated subpopulations and/or fluctuating chromosomal loci (Fig. 4b and Extended Data Fig. 10d, f). Similar analysis of fibroblast high-variance loci (199 X-linked and 127 autosomal) suggested perfect clonal maintenance of high-variance loci at both the X chromosome and autosomes (Fig. 4c and Extended Data Fig. 10f–h). Despite considerable methylation turnover rates, cells are capable of preserving a clear epigenetic distinction between alleles at these sites. We hypothesize that this can be achieved by propagating methylated and unmethylated patterns in a semi-faithful manner while relying on the larger-scale epigenomic context (for example, active/inactive X, active/inactive promoter, early/late replication) to correct errors and maintain the monoallelic state.

We studied the distributions of methylation patterns in clonal and matched polyclonal ES-cell populations to show that, in the pluripotent state, epigenomes are well organized and highly stable despite (or perhaps owing to) high DNA methylation turnover rates. ES cells show little if any evidence for direct transmission of clonal methylation memory. Such rapid epigenomic turnover processes can converge into a highly specific stationary methylation distribution within one or a few cell cycles, providing an explanation for the near perfect epigenomic reprogramming seen after induction of pluripotency genes⁸. This dynamic regime

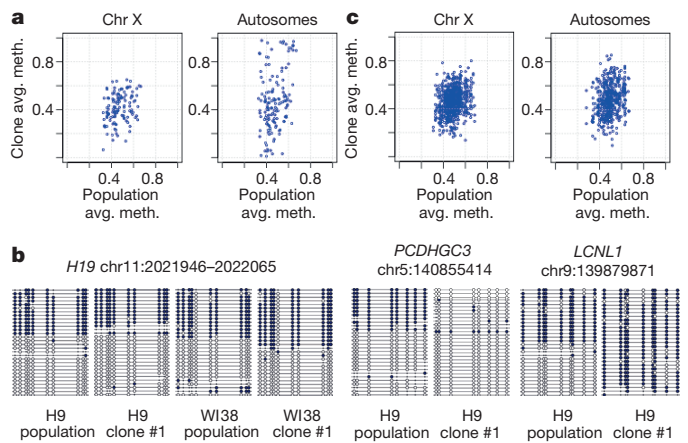


Figure 4 | Maintenance of monoallelic loci in stem cells and fibroblasts. **a**, Comparison of average methylation in clonal and polyclonal ES-cell populations, showing data for X-linked loci (H9 only) and autosomal loci (H9 and H1) with high variance (locus standard deviation >0.4) in the polyclonal population. avg. meth., average methylation; Chr, chromosome. **b**, Shown are methylation patterns observed for the *H19* imprinting control locus (stable biallelic), and the *PCDHGC3* and *LCNL1* loci (non-stable biallelic) **c**, Similar to **a** but for clonal and polyclonal fibroblast populations.

also suggests that the inherently transient methylation state of pluripotent cells *in vivo*⁶ may be unstable by design, and highlights the importance of characterizing the differences between such transient *in vivo* states and the long-term ES-cell models.

In contrast to ES cells, human fibroblasts, as well as three different human cancer cell lines and primary CD8⁺ mouse T cells show considerable clonal methylation memory. Turnover rates in clonal somatic populations are sufficiently low to allow phylogenetic reconstruction of the methylome of the founding single cell. This allows for detailed analysis of the methylation dynamics underlying the gradual loss of epigenomic homogeneity that is observed at markedly variable rates in different genomic contexts. We propose that a model combining clonal transmission of methylation patterns with slow but highly context-specific turnover rates can provide a unified explanation for the emergence of broad hypomethylated domains in late replicating regions^{27–30}, the accumulation of hypermethylation in polycomb-bound CpG islands^{19,20}, and the general age-coupled accumulation of methylation noise in many tissues. Although our quantitative turnover estimations are based on cell culture experiments, genome-wide methylation analysis in adult tissues suggests that *in vivo* dynamics converge to similar distributions and are thereby likely to be governed by similar principles.

More broadly, our analysis motivates further development of strategies for studying epigenomic organization as a dynamic property of populations of single cells. This may be essential for understanding the complex interactions between *trans*-acting factors, DNA methylation, chromatin modifications and chromosomal architectures, which contribute together and non-hierarchically to form and modify epigenetic memory in development and disease.

METHODS SUMMARY

UMI-bis reduced representation bisulphate sequencing was performed using a library of barcoded plasmids ligated to MspI fragmented DNA, and following an amended reduced representation bisulphate sequencing procedure. See Methods for additional details on UMI-bis and the cell growth protocols used. Processing of sequencing data and inference of pattern distributions and dynamics was performed by identifying groups of CpG sites with adequate common coverage (>20×) in sequencing libraries, inference of pattern frequencies per clonal or polyclonal population, and approximated sampling of the coalescence structures of such patterns (representing epialleles) within an exponentially growing population. This provided sufficient statistics for estimating rates of methylation gain and loss per CpG, with rate estimation becoming robust as more clones and/or more loci were pooled together.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 November 2013; accepted 7 May 2014.

Published online 13 July 2014.

- Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
- Cedar, H. & Bergman, Y. Programming of DNA methylation patterns. *Annu. Rev. Biochem.* **81**, 97–117 (2012).
- Hemberger, M., Dean, W. & Reik, W. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature Rev. Mol. Cell Biol.* **10**, 526–537 (2009).
- Mohn, F. & Schübeler, D. Genetics and epigenetics: stability and plasticity during cellular differentiation. *Trends Genet.* **25**, 129–136 (2009).
- Smallwood, S. A. *et al.* Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature Genet.* **43**, 811–814 (2011).
- Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
- Gifford, C. A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).
- Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65–70 (2013).
- Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–290 (2010).
- Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. & Cedar, H. Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proc. Natl Acad. Sci. USA* **79**, 61–65 (1982).
- Goll, M. G. & Bestor, T. H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74**, 481–514 (2005).
- Arand, J. *et al.* *In vivo* control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet.* **8**, e1002750 (2012).
- Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 (2013).
- Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Commun.* **2**, 241 (2011).
- Landan, G. *et al.* Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nature Genet.* **44**, 1207–1214 (2012).
- Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nature Methods* **9**, 72–74 (2012).
- Gu, H. *et al.* Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols* **6**, 468–481 (2011).
- Halligan, D. L. & Keightley, P. D. Spontaneous mutation accumulation studies in evolutionary genetics. *Annu. Rev. Ecol. Evol. Syst.* **40**, 151–172 (2009).
- Gal-Yam, E. N. *et al.* Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc. Natl Acad. Sci. USA* **105**, 12979–12984 (2008).
- Schlesinger, Y. *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. *Nature Genet.* **39**, 232–236 (2007).
- Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).
- Hansen, R. S. *et al.* Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc. Natl Acad. Sci. USA* **107**, 139–144 (2010).
- Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
- Lienert, F. *et al.* Identification of genetic elements that autonomously determine DNA methylation states. *Nature Genet.* **43**, 1091–1097 (2011).
- Kelly, T. K. *et al.* Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res.* **22**, 2497–2506 (2012).
- Deal, R. B., Henikoff, J. G. & Henikoff, S. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* **328**, 1161–1164 (2010).
- Hansen, K. D. *et al.* Increased methylation variation in epigenetic domains across cancer types. *Nature Genet.* **43**, 768–775 (2011).
- Berman, B. P. *et al.* Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nature Genet.* **44**, 40–46 (2012).
- Coolen, M. W. *et al.* Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. *Nature Cell Biol.* **12**, 235–246 (2010).
- Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Kenigsberg, E. Yaffe and the Tanay group for discussions. Research in the Tanay group was supported by the European Research Council (EVOEPIC), the EU BLUEPRINT project, the Israel Science Foundation (1050/12 and I-Core) the Israel Ministry of Science and the Helen and Martin Kimmel Award.

Author Contributions Z.S., Z.M. and A.T. designed the study. Z.S. and Z.M. performed the experiments. Z.S., N.M.C. and A.T. developed the algorithms. Z.S., Z.M. and A.T. analysed the data. G.L. and E.C. helped in developing the experimental protocol and analytical framework. Y.C.F. and E.A. generated ES-cell clones. S.R.Z. and N.F. generated CD8⁺ T-cell clones. Z.S., Z.M. and A.T. wrote the paper.

Author Information Raw data have been deposited in the Gene Expression Omnibus under accession number GSE53610. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.T. (amos.tanay@weizmann.ac.il).

Saturation editing of genomic regions by multiplex homology-directed repair

Gregory M. Findlay^{1*}, Evan A. Boyle^{1*}, Ronald J. Hause¹, Jason C. Klein¹ & Jay Shendure¹

Saturation mutagenesis^{1,2}—coupled to an appropriate biological assay—represents a fundamental means of achieving a high-resolution understanding of regulatory³ and protein-coding⁴ nucleic acid sequences of interest. However, mutagenized sequences introduced *in trans* on episomes or via random or “safe-harbour” integration fail to capture the native context of the endogenous chromosomal locus⁵. This shortcoming markedly limits the interpretability of the resulting measurements of mutational impact. Here, we couple CRISPR/Cas9 RNA-guided cleavage⁶ with multiplex homology-directed repair using a complex library of donor templates to demonstrate saturation editing of genomic regions. In exon 18 of *BRCA1*, we replace a six-base-pair (bp) genomic region with all possible hexamers, or the full exon with all possible single nucleotide variants (SNVs), and measure strong effects on transcript abundance attributable to nonsense-mediated decay and exonic splicing elements. We similarly perform saturation genome editing of a well-conserved coding region of an essential gene, *DBR1*, and measure relative effects on growth that correlate with functional impact. Measurement of the functional consequences of large numbers of mutations with saturation genome editing will potentially facilitate high-resolution functional dissection of both *cis*-regulatory elements and *trans*-acting factors, as well as the interpretation of variants of uncertain significance observed in clinical sequencing.

Functional consequences of genetic variants are best studied by manipulating the endogenous locus, which provides the native chromosomal context with respect to DNA sequence and epigenetic milieu, and for proteins, endogenous levels and patterns of expression⁷. Programmable endonucleases, for example, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) or clustered regularly interspaced short palindromic repeat (CRISPR)/Cas-based RNA-guided DNA endonucleases, enable direct genome editing with increasing practicality⁸. However, genome editing has primarily been applied to introduce single changes to one or a few genomic loci⁹, rather than many programmed changes to a single genomic locus.

We sought to leverage CRISPR/Cas9^{6,10,11} to introduce saturating sets of programmed edits to a specific locus via multiplex homology-directed repair (HDR). We first targeted six bases of a *BRCA1* exon¹². We cloned an HDR library containing random hexamers substituted at positions +5 to +10 of *BRCA1* exon 18 and fixed, nonsynonymous changes at positions +17 to +23 (as a ‘handle’ for selective PCR and to prevent re-cutting¹³ by destroying the protospacer adjacent motif (PAM)) (Fig. 1a; Supplementary Table 1). We co-transfected pCas9-sgBRCA1x18 and the HDR library into ~800,000 HEK293T cells, achieving 3.33% HDR efficiency. We performed two independent transfections with the same HDR library (‘biological replicates’ 1, 2), and cells were split on day 3 (‘D3 replicates’ a, b).

We prepared genomic DNA (gDNA) and complementary DNA (cDNA) from bulk cells on D5. PCR reactions were primed on the ‘handle’ uniquely present within successfully edited genomes. Amplification was observed in HDR library/pCas9-sgBRCA1x18-transfected samples, but not in HDR library-only controls. Amplicons derived from gDNA

and cDNA were deeply sequenced (Fig. 1a). The relative abundances of hexamers within replicates and the correlation between the HDR library and edited gDNA were consistent with limited ‘bottlenecking’ during transfection and minimal influence of hexamer identity on HDR efficiency (Extended Data Figs 1 and 2).

We estimated the effect of introducing each hexamer to these genomic coordinates on transcript abundance by calculating enrichment scores (cDNA divided by gDNA counts, calibrated to wild type). These enrichment scores were well correlated between biological replicates (Fig. 1b, 1a vs 2a: $R = 0.659$) and between D3 replicates (Extended Data Fig. 2c; 1a vs 1b: $R = 0.662$). When we pooled read counts from D3 replicates, correlation between biological replicates improved (Extended Data Fig. 2d; 1 vs 2: $R = 0.706$).

To maximize precision (see Supplementary Note 1 for discussion of reproducibility), we merged data across all four replicates for 4,048 hexamers (Fig. 1c; Supplementary Table 2). Several results support the biological validity of the resulting enrichment scores. First, as anticipated by nonsense-mediated decay (NMD), hexamers introducing stop codons were associated with markedly reduced mRNA levels (Fig. 1c; Wilcoxon rank sum test (WRST) $P = 9.7 \times 10^{-84}$; median for nonsense hexamers 12-fold below overall median). Second, previous studies measured hexamer influence on splicing at analogous coordinates of different exons via a plasmid minigene assay¹⁴. Despite these contextual differences, the strongest exonic splicing silencers (ESSs) (bottom 2% in ref. 14) scored ninefold below median (Fig. 1c; WRST $P = 2.0 \times 10^{-24}$), the strongest exonic splicing enhancers (ESEs) (top 2% in ref. 14) scored 1.5-fold above median (Fig. 1c; WRST $P = 2.4 \times 10^{-11}$), and the complete data sets correlated reasonably well (Extended Data Fig. 3a; $\rho = 0.524$). We also observed correlation between G+C content and enrichment scores (Extended Data Fig. 3b), strongest for bases most proximal to the splice junction, consistent with a posited role for G+C content in the stability of splicing structures¹⁵ (although reverse transcription bias is a potential confounder).

We next sought to assay the effects of SNVs across the full 78 base pairs of *BRCA1* exon 18 (Extended Data Fig. 4). We cloned three HDR libraries with selective PCR sites in either the 5′ or 3′ region and 3% doping¹⁶ (97(WT):1:1:1) in the other half of the exon (L: 5′ degeneracy, 3′ nonsynonymous selective PCR site; R: 3′ degeneracy, 5′ nonsynonymous selective PCR site; R2: 3′ degeneracy, 5′ synonymous selective PCR site) (Supplementary Table 1). Five days post-transfection with pCas9-sgBRCA1x18 (1.02–1.29% HDR efficiency), we selectively amplified and deeply sequenced gDNA and cDNA.

Using data from all edited exons with ≥ 1 mutation and ≥ 10 gDNA counts, we estimated effect sizes of all possible SNVs using a weighted linear model. Estimated effect sizes were reproducible ($R = 0.846$ (R), 0.853 (R2), and 0.686 (L); Fig. 2a, Extended Data Figs 5 and 6, Supplementary Table 3). Effect sizes for the same SNVs interrogated with different selective PCR strategies (R vs R2) were also well correlated ($R = 0.847$; Fig. 2b).

The estimated effect sizes reflect empirically measured changes in transcript abundance resulting from programmed edits (Fig. 2c). As expected with NMD, nonsense mutations reduced transcript abundance

¹Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA.

*These authors contributed equally to this work.

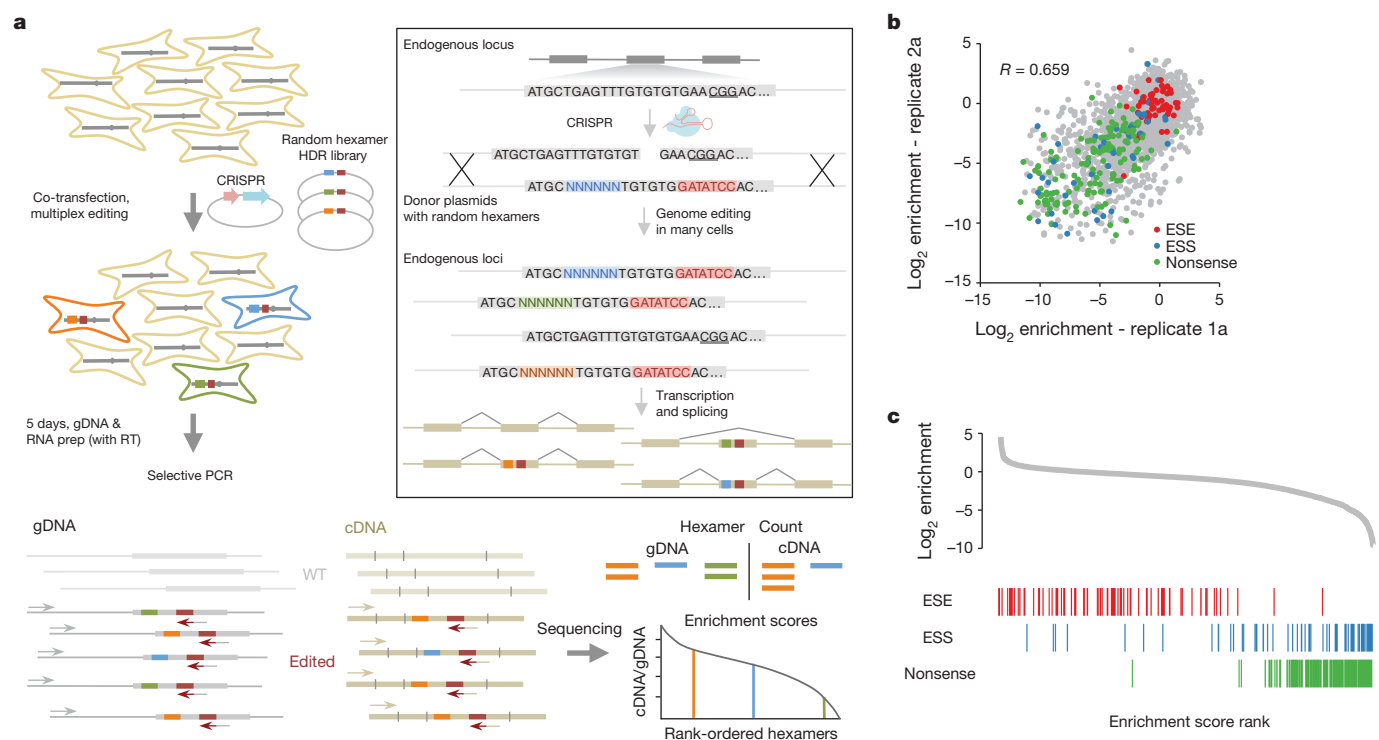


Figure 1 | Saturation genome editing and multiplex functional analysis of a hexamer region influencing *BRCA1* splicing. **a**, Experimental schematic. Cultured cells were co-transfected with a single Cas9-sgRNA construct (CRISPR) and a complex homology-directed repair (HDR) library containing an edited exon that harbours a random hexamer (blue, green, orange) and a fixed selective PCR site (red). CRISPR-induced cutting stimulated homologous recombination with the HDR library, inserting mutant exons into the genomes of many cells. At five days post-transfection, cells were harvested for gDNA

and RNA. After reverse transcription, selective PCR was performed followed by sequencing of gDNA- and cDNA-derived amplicons. Hexamer enrichment scores were calculated by dividing cDNA counts by gDNA counts.

b, Correlation of enrichment scores between biological replicates for hexamers observed in each experiment with positions of previously identified¹⁴ exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs) and stop codons indicated. **c**, Rank-ordered plot of enrichment scores with positions of ESEs, ESSs and stop codons indicated.

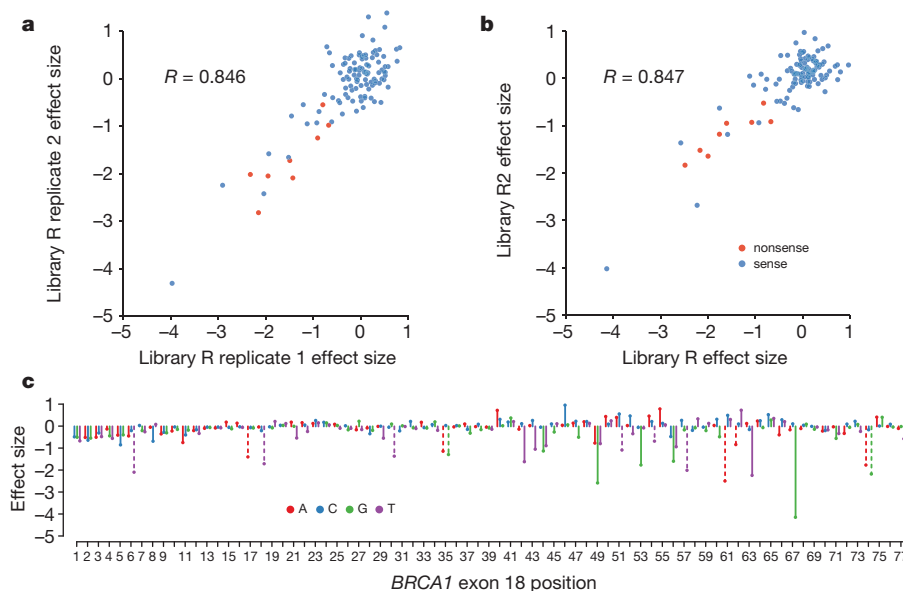


Figure 2 | Multiplex homology-directed repair reveals effects of single nucleotide variants on transcript abundance. Three separate HDR libraries (R, R2, and L) containing a 3% mutation rate (97% WT, 1% each non-WT base) in either half of *BRCA1* exon 18 were introduced to the genome via co-transfection with pCas9-sgBRCA1x18. Enrichment scores were calculated for each haplotype observed at least 10 times in gDNA sequencing, and effect sizes of SNVs were determined by weighted linear regression modelling. 'Sense' includes both missense and synonymous SNVs. **a**, Effect sizes calculated from replicate transfections of HDR library R, consisting of a 3% per-nucleotide

mutation rate in the 3'-most 39 bases and the same selective PCR site used in Fig. 1, were highly correlated ($R = 0.846$). **b**, Library R2 harboured a selective PCR site composed of 5 synonymous changes, none of which are present in library R. When effect sizes derived from experiments with library R2 were plotted against those from library R, there was a strong correlation ($R = 0.847$), indicating reproducibility and demonstrating that differences between selective PCR sites did not strongly influence scores. **c**, Effect sizes for SNVs across the exon are displayed. Data sets from libraries R and L were combined to span the entire exon. Dashed lines represent SNVs that introduce nonsense codons.

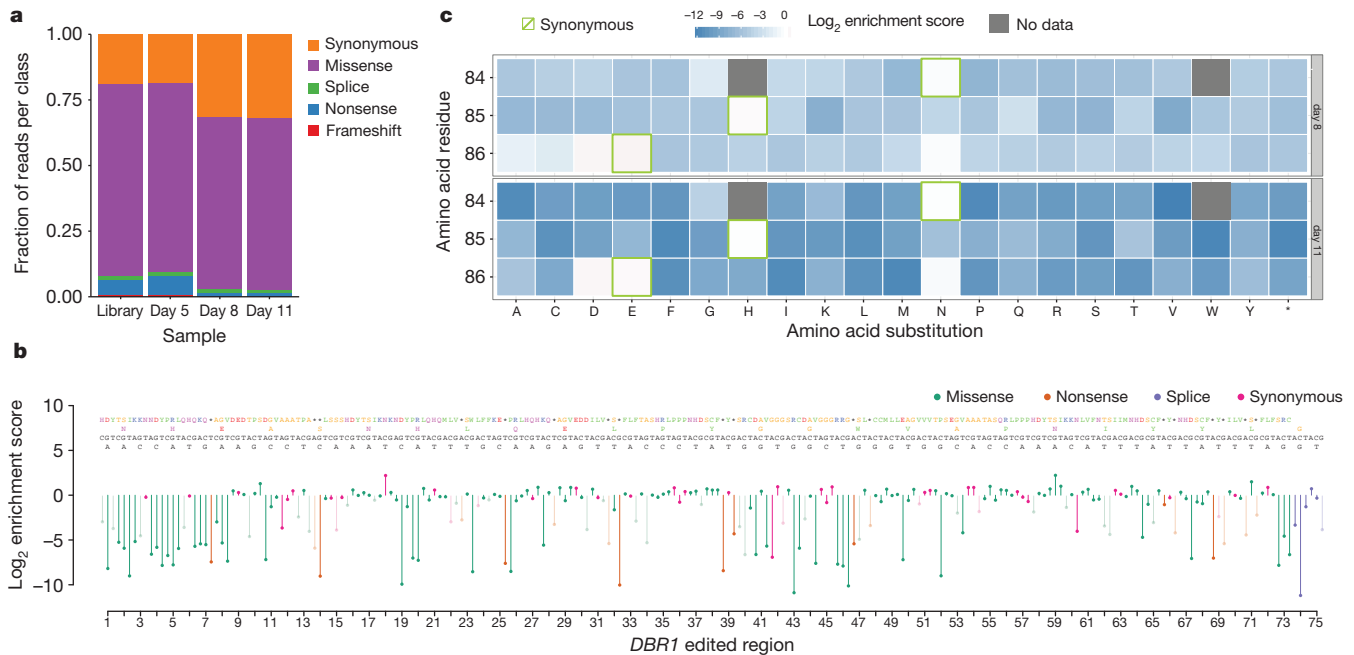


Figure 3 | Saturation genome editing and multiplex functional analysis at an essential gene, *DBR1*, in Hap1 cells. An HDR library targeting a highly conserved region of *DBR1* exon 2 was used with pCas9-EGFP-sgDbr1x2 to introduce point mutations across 75 bp and all possible codon substitutions at three residues believed to participate at the enzyme's active site. **a**, Sequencing of gDNA from the HDR library and populations of edited cells at D5, D8 and D11 reveals selection for synonymous mutations, and depletion of frameshift, nonsense and missense variants. **b**, Mean D11 enrichment scores are plotted as line segments for SNVs in the 3'-most 73 bases of exon 2 and two bases of intron 2. Above the enrichment scores in ascending order are the WT nucleotide at each position, each 1-bp genome edit, the wild-type amino

(WRST $P = 1.4 \times 10^{-203}$; 5.6-fold below median). Additionally, several missense and synonymous SNVs reproducibly resulted in large reductions in transcript abundance, and SNV effect sizes correlated with a predictive model for exonic variants that disrupt splicing¹⁷ ($\rho = 0.322$; Extended Data Fig. 7a). Because HDR with library L does not destroy the PAM, we calculated enrichment scores for indels from non-homologous end-joining (NHEJ). As expected with NMD, only frameshifting indels were associated with large depletions (Extended Data Fig. 7b, c).

To further demonstrate this method, we targeted a well-conserved region of *DBR1*, the RNA lariat debranching enzyme, which scored highly in a genome-wide screen for essentiality¹⁸ (Extended Data Fig. 8). We used array-synthesized oligonucleotides to program a *DBR1* HDR library to include the wild-type sequence and every possible SNV across 75 bp (73 3'-most bases of exon 2 and first two bases of intron 2), and also all 63 possible codon substitutions at three residues (388 genome edits were programmed; single base deletions were abundant from synthesis errors). The HDR library also introduced two fixed synonymous changes (to disrupt the PAM and prevent re-cutting¹³) and a selective PCR site in intron 2.

An optimized single-guide RNA (sgRNA) sequence^{19,20} was cloned into a bicistronic sgRNA/Cas9-2A-EGFP vector (pCas9-EGFP-sgDbr1x2). Five million haploid human cells²¹ (Hap1) were co-transfected with the *DBR1* HDR library and pCas9-EGFP-sgDbr1x2. On D2, ~250,000 enhanced green fluorescent protein-positive (EGFP⁺) cells were sorted by fluorescence-activated cell sorting (FACS) and further cultured, taking samples on D5, D8 and D11 (1.14% HDR efficiency, estimated on D8). Following gDNA isolation and selective PCR, deep sequencing was performed to quantify the relative abundance of edited haplotypes in each sample.

We first examined the relative proportions of mutation classes at each time point (Fig. 3a). The strong enrichment of synonymous mutations

acid (AA), and the AA derived from each genome edit (asterisk indicates a stop codon). Segment colour indicates mutation type, faded segments indicate discordant effects between replicates, and AAs are coloured according to the Lesk colour scheme (orange, small nonpolar; green, hydrophobic; magenta, polar; red, negatively-charged; blue, positively charged). The first nine bases shown correspond to the active site residues. **c**, D8 (top) and D11 (bottom) amino acid level enrichment scores were calculated for active site residues N84, H85, E86 after excluding discordant observations between replicates (Extended Data Fig. 10c). On both D8 and D11 we observe strong selective effects and tolerance of only synonymous (green boxes) and a few missense variants.

and depletion of nonsense and frameshifting mutations over time indicated that selection was acting on edited cells in culture, consistent with *DBR1* essentiality. We calculated enrichment scores (D8 or D11 counts divided by D5 counts) for 365 of the 388 (94%) programmed edits and 12 single base deletions (the subset with relative abundance $> 5 \times 10^{-5}$ on D5) (Fig. 3b; Extended Data Fig. 9; Supplementary Table 4). Enrichment scores strongly correlated with functional consequence. The median enrichment score for synonymous edits was nearly identical to wild-type (1.006-fold lower), but 73-fold lower for missense edits ($P = 1.7 \times 10^{-8}$; WRST against synonymous edits), 207-fold lower for nonsense edits ($P = 1.9 \times 10^{-9}$), and 211-fold lower for frameshifting single base deletion edits ($P = 1.5 \times 10^{-8}$). Furthermore, enrichment scores for SNVs were inversely correlated with metrics of predicted deleteriousness like CADD²² ($\rho = -0.295$; $P = 1.2 \times 10^{-5}$; Extended Data Fig. 10a, b). Residues N84, H85 and E86 of *DBR1* were edited to all 63 possible non-wild-type codons. Consistent with their predicted role in the active site of an essential enzyme²³, only synonymous mutations and a few missense substitutions were tolerated (Fig. 3c).

Amino acid level enrichment scores were well correlated between D11 biological replicates ($R = 0.752$; $P = 2.6 \times 10^{-40}$; Extended Data Fig. 10c), and were bimodally distributed in each replicate, allowing broad classification of changes as tolerated or deleterious. The small proportion of discordantly classified variants might be explained by Hap1 reversion to diploidy or off-target effects, highlighting the importance of biological replicates for this experimental design (Supplementary Note 1). Notably, there were no reproducibly tolerated nonsense or frameshifting edits. Overall, these data support the conclusion that our empirically derived enrichment scores reflect true biological effects of specific genomic point mutations within *DBR1*.

We demonstrate that it is feasible to generate and functionally analyse hundreds to thousands of programmed genome edits at a single

locus in a single experiment. We emphasize three major limitations of the method as it stands. First, we only introduced programmed edits to the immediate vicinity of coordinates targeted by the endonuclease (Extended Data Figs 5a and 9a), and the narrow window associated with HDR mechanisms in mammalian cells²⁴ may fundamentally limit the size of the region that can be subjected to multiplex editing in one experiment. Saturation genome editing of a full gene—for example, to measure functional consequences of all possible variants of uncertain significance—will require multiple experiments tiling along its exons.

Second, only a small proportion of cells were successfully edited in each experiment, bottlenecking complexity, limiting reproducibility (Supplementary Note 1), and necessitating the selective PCR site. Looking forward, a variety of techniques, for example, transient hypothermia²⁵ or oligonucleotide-based HDR²⁶, may improve editing efficiency. Consistent with this, we note that ZFNs and TALENs have demonstrated efficiencies up to 50% in some studies^{27,28}. Also, although the low editing efficiency necessitated using haploid cells for *DBRI* mutagenesis, this could potentially have been performed in diploid cells by knocking out one allele via NHEJ and then knocking in the HDR library to the other allele.

Finally, the development of functional assays that are biologically relevant and technically viable remains a challenge. Here, we exploited strategies that directly linked genotype to phenotype—for example, targeted RNA sequencing to measure transcript abundance or targeted DNA sequencing to measure reduced cellular fitness. Analogous approaches can be taken in other contexts—for example, targeted chromatin immunoprecipitation-sequencing (ChIP-seq) of co-activators to assay enhancers, increased cellular growth rate to assay cancer drivers or drug resistance²⁹, or FACS-based phenotypic sorting for cellular assays more generally³⁰ (Supplementary Note 2).

There is a strong demand for techniques that accurately and scalably measure mutational consequences, and a dearth of experimental data measuring distributions of effect sizes or corresponding to direct manipulation of the genome. By multiplexing both the introduction and assaying of mutations in their native context, we anticipate that saturation genome editing will accelerate our ability to measure and interpret the functional consequences of genetic variation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 May; accepted 18 July 2014.

Published online 20 August 2014.

- Myers, R. M., Tilly, K. & Maniatis, T. Fine structure genetic analysis of a beta-globin promoter. *Science* **232**, 613–618 (1986).
- Cunningham, B. C. & Wells, J. A. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244**, 1081–1085 (1989).
- Patwardhan, R. P. *et al.* High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nature Biotechnol.* **27**, 1173–1175 (2009).
- Fowler, D. M. *et al.* High-resolution mapping of protein sequence-function relationships. *Nature Methods* **7**, 741–746 (2010).
- Botstein, D. & Shortle, D. Strategies and applications of *in vitro* mutagenesis. *Science* **229**, 1193–1201 (1985).
- Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Gibson, T. J., Seiler, M. & Veitia, R. A. The transience of transient overexpression. *Nature Methods* **10**, 715–721 (2013).

- Gaj, T., Gersbach, C. A. & Barbas, C. F. III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
- Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
- Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
- Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Mazoyer, S. *et al.* A BRCA1 nonsense mutation causes exon skipping. *Am. J. Hum. Genet.* **62**, 713–715 (1998).
- Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnol.* **32**, 347–355 (2014).
- Ke, S. *et al.* Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.* **21**, 1360–1374 (2011).
- Zhang, J., Kuo, C. C. & Chen, L. GC content around splice sites affects splicing through pre-mRNA secondary structures. *BMC Genomics* **12**, 90 (2011).
- Patwardhan, R. P. *et al.* Massively parallel functional dissection of mammalian enhancers *in vivo*. *Nature Biotechnol.* **30**, 265–270 (2012).
- Mort, M. *et al.* MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing. *Genome Biol.* **15**, R19 (2014).
- Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
- Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnol.* **31**, 827–832 (2013).
- Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281–2308 (2013).
- Carette, J. E. *et al.* Haploid genetic screens in human cells identify host factors used by pathogens. *Science* **326**, 1231–1235 (2009).
- Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genet.* **46**, 310–315 (2014).
- Khalid, M. F., Damha, M. J., Shuman, S. & Schwer, B. Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Res.* **33**, 6349–6360 (2005).
- Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J. A. & Jasin, M. Gene conversion tracts from double-strand break repair in mammalian cells. *Mol. Cell Biol.* **18**, 93–101 (1998).
- Doyon, Y. *et al.* Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. *Nature Methods* **7**, 459–460 (2010).
- Chen, F. *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature Methods* **8**, 753–755 (2011).
- Reyon, D. *et al.* FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnol.* **30**, 460–465 (2012).
- Carroll, D. Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* **83**, 409–439 (2014).
- Smurnyy, Y. *et al.* DNA sequencing and CRISPR-Cas9 gene editing for target validation in mammalian cells. *Nature Chem. Biol.* **10**, 623–625 (2014).
- Kinney, J. B., Murugan, A., Callan, C. G., Jr & Cox, E. C. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proc. Natl Acad. Sci. USA* **107**, 9158–9163 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank F. Zhang and his laboratory for the CRISPR/Cas9 backbone constructs used in this study and G. Church and his laboratory for providing reagents used to establish CRISPR/Cas9 editing techniques in our lab. We also thank members of the Shendure laboratory for helpful discussions and D. Prunkard for assistance with FACS. This work was supported by the National Institutes of Health (DP1HG007811 to J.S.) and the UW Medical Scientist Training Program (G.M.F. and J.K.).

Author Contributions The project was conceived and designed by G.M.F. and J.S. G.M.F. and E.A.B. performed experiments. E.A.B. and R.J.H. performed data analysis and generated data figures. G.M.F. generated schematic figures. G.M.F., E.A.B., R.J.H. and J.S. wrote the manuscript. J.C.K. assisted G.M.F. to establish genome editing techniques in the laboratory.

Author Information Sequence data used for this analysis are available in SRA under accession number SRP044126. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.S. (shendure@uw.edu) or G.M.F. (gf2@uw.edu).

High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875

Ankita Srivastava^{1*}, Jason Yano^{1*}, Yoshihiko Hirozane², Georgia Kefala¹, Franz Gruswitz^{1†}, Gyorgy Snell¹, Weston Lane¹, Anthony Ivetac¹, Kathleen Aertgeerts^{1†}, Jasmine Nguyen^{1†}, Andy Jennings¹ & Kengo Okada²

Human GPR40 receptor (hGPR40), also known as free fatty-acid receptor 1 (FFAR1), is a G-protein-coupled receptor that binds long-chain free fatty acids to enhance glucose-dependent insulin secretion¹. Novel treatments for type-2 diabetes mellitus² are therefore possible by targeting hGPR40 with partial or full agonists. TAK-875, or fasiglifam, is an orally available, potent and selective partial agonist³ of hGPR40 receptor, which reached phase III clinical trials for the potential treatment of type-2 diabetes mellitus⁴. Data from clinical studies indicate that TAK-875, which is an ago-allosteric modulator of hGPR40 (ref. 3), demonstrates improved glycaemic control and low hypoglycaemic risk in diabetic patients⁵. Here we report the crystal structure of hGPR40 receptor bound to TAK-875 at 2.3 Å resolution. The co-complex structure reveals a unique binding mode of TAK-875 and suggests that entry to the non-canonical binding pocket most probably occurs via the lipid bilayer. The atomic details of the extensive charge network in the ligand binding pocket reveal additional interactions not identified in previous studies and contribute to a clear understanding of TAK-875 binding to the receptor. The hGPR40–TAK-875 structure also provides insights into the plausible binding of multiple ligands to the receptor, which has been observed in radioligand binding⁶ and Ca²⁺ influx assay studies³. Comparison of the

transmembrane helix architecture with other G-protein-coupled receptors suggests that the crystallized TAK-875-bound hGPR40 complex is in an inactive-like state.

Human GPR40 binds free fatty acids and is primarily localized in pancreatic β-cells and intestinal enteroendocrine cells^{1,2,7}. Natural substrates of hGPR40 include saturated C12–C16 or unsaturated C18–C20 free fatty acids, which act as full agonists of hGPR40 to mediate glucose-stimulated insulin secretion from pancreatic β-cells and glucagon-like peptide 1 (GLP-1) secretion from intestinal cells^{7,8}. The hGPR40 receptor mainly couples with the Gαq G-protein subtype upon agonist stimulation to mediate downstream signalling by increasing intracellular calcium concentration, [Ca²⁺]_i (ref. 9). TAK-875 is a hGPR40 partial agonist³ which stimulates glucose-dependent insulin secretion, demonstrating a unique mode of action for the treatment of type-2 diabetes mellitus⁵.

To understand the mechanism of action and the interactions of TAK-875 with hGPR40 better, we crystallized the hGPR40–TAK-875 co-complex in lipidic cubic phase and determined the structure at 2.3 Å resolution by X-ray crystallography. Four point mutations (Leu-42^{2,40}Ala, Phe-88^{3,34}Ala, Gly-103^{3,49}Ala and Tyr-202^{5,58}Phe) were made to increase expression levels and thermal stability, and a T4 lysozyme (T4L) fusion protein was inserted into the third intracellular loop (ICL3) to support

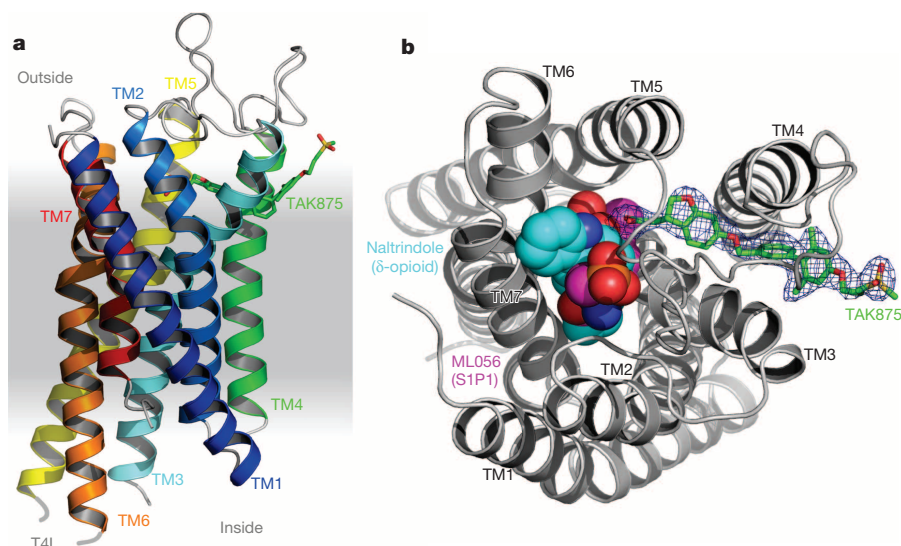


Figure 1 | Structure of hGPR40 in complex with the partial agonist TAK-875. **a**, The conserved seven transmembrane helical bundle fold is represented as rainbow ribbons while TAK-875, bound between transmembrane helix (TM) 3 and transmembrane helix 4, is shown as a stick model (carbon, green; sulphur, yellow; oxygen, red). T4L, inserted between transmembrane helix 5 and transmembrane helix 6, is illustrated schematically. **b**, Extracellular view of

hGPR40. The $1F_o - F_c$ omit map of the ligand was calculated in the absence of TAK-875 and is shown as a blue mesh contoured at 3σ . To highlight the unique binding mode of TAK-875, the receptor portions of the S1P1-ML056 (3V2Y) and δ-opioid-naltrindole (4EJ4) complexes were aligned to hGPR40 and their respective ligands are shown as spheres (ML056, magenta; naltrindole, cyan).

¹Department of Structural Biology and Core Sciences & Technology, Takeda California, 10410 Science Center Drive, San Diego, California 92121, USA. ²Biomolecular Research Laboratories, Pharmaceutical Research Division, Takeda Pharmaceutical Company Ltd, 26-1, Muraoka-Higashi 2-chome, Fujisawa, Kanagawa 251-8555, Japan. [†]Present addresses: Beryllium, Membrane Protein Sciences, 7869 NE Day Road West, Bainbridge Island, Washington 98110, USA (F.G.); Dart Neuroscience, 12278 Scripps Summit Drive, San Diego, California 92131, USA (K.A. and J.N.).

*These authors contributed equally to this work.

crystal growth (Extended Data Fig. 1a–c). The binding affinity of TAK-875 to the receptor is not significantly affected by these mutations or T4L insertion with the dissociation constant (K_d) in the approximate range 12–16 nM (Extended Data Fig. 2). On the basis of a Ca^{2+} flux assay, a construct with the four point mutations and native ICL3 revealed reduced potency by more than 1000-fold compared with the wild-type (WT) receptor¹⁰. This suggests that the thermostabilizing mutations do not affect the TAK-875 binding pocket but trap the receptor in an agonist-bound inactive-like conformation by restraining the conformational changes required for G-protein coupling and receptor activation. Similar observations have also been reported for other agonist-bound G-protein-coupled receptor (GPCR) structures containing thermostabilizing mutations, such as NTSR1 (ref. 11) and A_{2A} (ref. 12).

The seven-transmembrane helices bundle, typical of GPCR structures, remains conserved in the hGPR40 structure (Fig. 1a). The hGPR40–TAK-875 co-complex structure reveals that TAK-875 uniquely binds to a non-canonical site created between transmembrane helices 3–5 and the extracellular loop 2 (ECL2) regions (Fig. 1a, b and Extended Data Fig. 3), closer to the exterior membrane surface than most other GPCR ligands. This site is roughly perpendicular to the canonical orthosteric site observed between transmembrane helix 3 and transmembrane helices 5–7 in other GPCR structures, such as S1P1 (ref. 13) and δ -opioid¹⁴ (Fig. 1b). TAK-875 has been characterized as an ago-allosteric modulator³, which suggests that this unique binding site represents an allosteric binding pocket in the receptor. ECL2 serves as the roof of the TAK-875 binding cavity with Glu 172^{ECL2} hydrogen bonded to Arg 258^{7,35}. Moreover, Trp 174^{ECL2} is oriented nearly perpendicularly to the plane of the dihydroisobenzofuran ring of TAK-875 where it forms an edge-on interaction (Fig. 2a, b).

Despite relatively low sequence identity, hGPR40 shows structural similarities with peptide-binding GPCRs such as NTSR1, PAR1, CXCR4 and opioid receptors^{11,15–17}. Notably, they all have a conserved hairpin loop motif in the extracellular loop 2 (ECL2 hairpin) and a disulphide bond formed between transmembrane helix 3 (Cys 79^{3,25} in hGPR40) and at the carboxy (C)-terminal portion of the ECL2 loop (Cys 170^{ECL2} in hGPR40) (Fig. 2b and Extended Data Fig. 4). In the peptide-binding and opioid GPCRs the ECL2 hairpin has a distinctive β -sheet reaching between helices 4 and 5, while in hGPR40 it is a shorter β -sheet-like region (marked in cyan in Fig. 2b and Extended Data Fig. 4). This β -sheet-like region in hGPR40 has low B-factors (Extended Data Fig. 5a–c), reflecting the low mobility of this region. This limits the overall flexibility of the adjacent portion of ECL2 between Leu 171 and Asp 175, which forms a cap over the canonical binding site. In addition, ECL2 is also stabilized by an extensive network of interactions with the transmembrane helices and within the loop itself, which further limits the mobility of ECL2 (Extended Data Fig. 5a–c). In all of the other peptide/opioid receptors, the amino (N)-terminal region of the ECL2 hairpin loop is directly connected to transmembrane helix 4, while in hGPR40 there are an additional 13 residues (Pro 147 to Gly 159, marked in pink in Fig. 2b and Extended Data Fig. 4), which form a separate auxiliary loop between the β -sheet-like region and transmembrane 4. The relatively low B-factors of the hGPR40 ECL2 (Extended Data Fig. 5a–c) indicate low flexibility of this region, with the only exception being the tip of the auxiliary loop (Asp 152–Asn 155). The stem of the auxiliary loop Ala 146 and Pro 147 on one side and Leu 158 on the other is stabilized by the conserved transmembrane helix 4 and hairpin loop, respectively (Extended Data Fig. 5a–c). Similar to S1P1 receptor and in contrast to the open cavity seen in the δ -opioid receptor, the entire hGPR40 ECL2 functions as a cap-like structure, which covers the binding pocket from the central extracellular region (Figs 1b and 2b and Extended Data Fig. 3).

The structural features of ECL2 and high lipophilicity of TAK-875 (which shows approximately 380-fold preference for lipophilic phase over aqueous phase) suggest that TAK-875 most probably enters the binding pocket of hGPR40 through the lipid bilayer, which is also supported by the protrusion of the 1,3-dimethyl-5-(3-(methylsulfonyl)propoxy) benzene group between transmembrane helix 3 and transmembrane helix 4 of hGPR40 (Fig. 2a, b), but it does not rule out the possibility that TAK-875

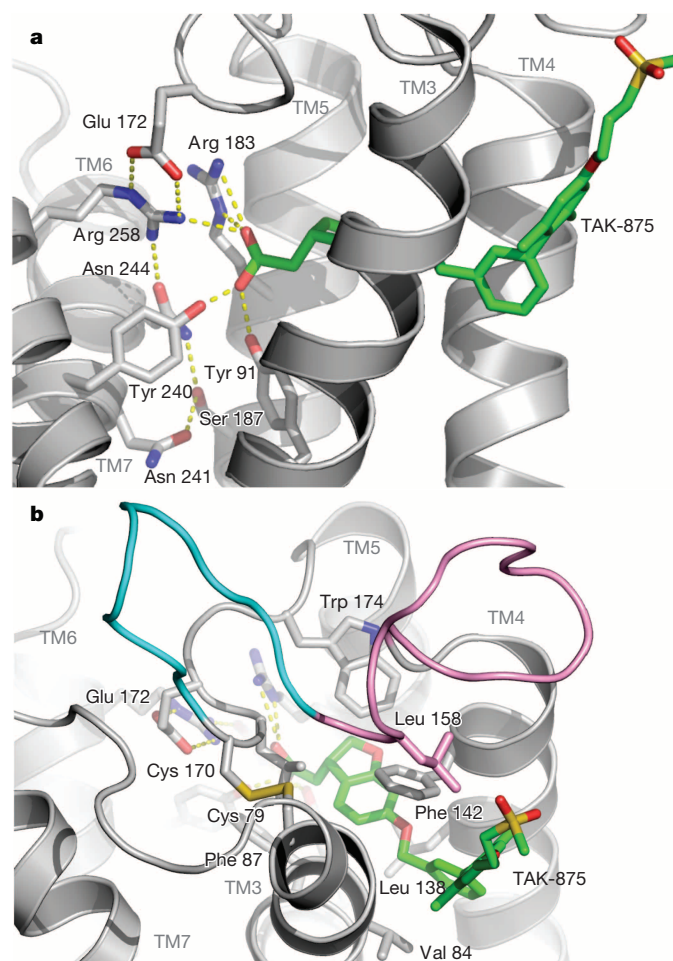


Figure 2 | Binding interactions of TAK-875. **a**, Interaction of the carboxylate of TAK-875 with the charge cluster as seen in a lateral view. The carboxylate moiety is highly coordinated by several key residues. hGPR40 is represented as grey ribbon and TAK-875 as sticks coloured by atom type, with carbon in green, sulphur in yellow and oxygen in red. The side chains are represented as sticks and coloured by atom type, with carbon in grey. Hydrogen bonds are represented by yellow dashes. **b**, An extracellular view of TAK-875 binding, showing the β -sheet-like region of ECL2 (coloured in cyan) and the disulphide between Cys 79^{3,25} and Cys 170^{ECL2}. These two features are conserved in peptide-binding receptors and hGPR40. The auxiliary loop formed between Pro 147 and Gly 159 is shown in pink as part of the ECL2. TAK-875 entry is positioned between transmembrane 3 and transmembrane 4.

can enter the binding pocket through lipid bilayer interface from the extracellular side between transmembrane helix 3 and transmembrane helix 4.

The entry of hydrophobic substrates to GPCRs through the lipid bilayer has been previously proposed for the retinal loading of opsin¹⁸ and for the entry of AM-841 and sn-2-arachidonylglycerol into the cannabinoid 2 receptor^{19,20}. Lipid entry for S1P1 (ref. 13) is suggested but from an unrelated trajectory rotated by 120° in the plane of the membrane compared with hGPR40 (Extended Data Fig. 3).

Like the PAR1–vorapaxar co-complex structure (Protein Data Bank accession number 3VW7), TAK-875 is located between two transmembrane helices. In the case of PAR1, the ethyl carbamate tail of vorapaxar is situated between transmembrane helix 6 and transmembrane helix 7, while in hGPR40 the linker connecting the dihydrobenzofuran to the benzyl ring is accommodated between transmembrane helix 3 and transmembrane helix 4. This binding mode of TAK-875 aligns with the structure–activity relationship determined in ref. 4, which identified that the 4'-position of the terminal biphenyl ring could tolerate a variety of modifications of differing length and polarity with minimal effect on

potency and is thus well suited to modify other properties of the compound. The sulphonate group of TAK-875 serves to reduce lipophilicity and to improve absorption, distribution, metabolism, excretion and toxicology (ADME-tox) properties and is not required for the agonistic activity⁴²¹.

The interaction between the carboxylate moiety of agonist ligands with Arg 183^{5,39} and Arg 258^{7,35}, as observed for TAK-875 (Fig. 2a and Extended Data Fig. 6), has precedent in a number of mutational, modelling and computational studies^{22,23}. In addition to the arginine residues, our crystal structure reveals the stabilization of the carboxylate moiety by two tyrosine residues, Tyr 91^{3,37} and Tyr 240^{6,51}, which have not been previously reported (Fig. 2a and Extended Data Fig. 6). Instead, the interactions of the tyrosine residues with the carboxylate moiety²² were predicted to be aromatic/hydrophobic, rather than the direct interactions observed in the crystal structure. The presence of the multiple arginine and tyrosine residues interacting with the carboxylate suggests a possible redundancy in the charge network, which is supported by the observation that certain carboxylate containing ligands are not affected to the same degree upon single mutation of these arginine residues. For example, the half-maximum effective concentration values of docosahexaenoic acid and AM1638 are only modestly affected by mutation of Arg 183^{5,39} or Arg 258^{7,35}. Similarly, phenylalanine mutants of Tyr 91^{3,37} or Tyr 240^{6,51} were shown to bind the agonists linoleic acid and GW9508 with near WT affinity^{22,23}. The carboxylate moiety of TAK-875 participates in a complex charge network that also involves Glu 172^{ECL2}, Ser 187^{5,43}, Asn 241^{6,53} and Asn 244^{6,55} (Fig. 2b and Extended Data Fig. 6). Mutational studies based on molecular modelling implied that this charge network keeps the receptor in a ligand-free, inactive state, and, upon agonist binding, the residues engaged in this charge network are disrupted²². Interestingly, in the hGPR40–TAK-875 partial agonist co-complex, the charge network (Fig. 2b and Extended Data Fig. 6) coordinates the carboxylate and remains intact.

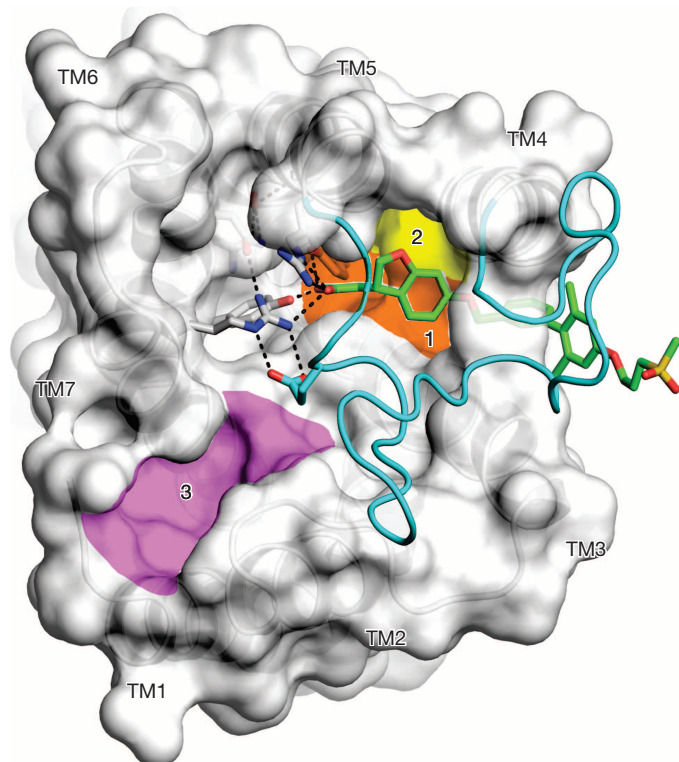


Figure 3 | Surface representation of hGPR40 showing multiple putative binding sites. Site 1 (orange) is the TAK-875 binding site observed in the hGPR40 crystal structure. The alternative proposed binding site, site 2 (yellow), would allow ligands to pass between transmembrane 4 and transmembrane 5, suggesting entry through the membrane. Site 3 (pink) is positioned between transmembrane 1 and transmembrane 7, suggesting extracellular entry. ECL2 is shown in cyan and the surface for ECL2 was removed for clarity.

Site-directed mutagenesis of the key binding-pocket residues Tyr-91^{3,37} Ala, Arg-183^{5,39} Ala, Asn-244^{6,55} Ala and Arg-258^{7,35} Ala of the WT hGPR40 and thermostabilized hGPR40 receptor had detrimental effects on the binding affinity (K_d) of TAK-875, while Tyr-240^{6,51} Ala reduced the binding affinity by approximately eightfold (Extended Data Figs 7 and 8 and Extended Data Table 1). Similarly in Ca^{2+} influx assays, the effect of Arg-183^{5,39} Ala, Arg-258^{7,35} Ala, Tyr-91^{3,37} Ala and Asn-244^{6,55} Ala single point mutations on WT hGPR40 drastically affected the TAK-875 potency³. Taken together, the mutagenesis studies confirm the critical role of these residues in TAK-875 binding and verify that the interactions observed between TAK-875 and the thermostabilized hGPR40 reflect the physiological binding mode of TAK-875.

Prior radioligand binding studies identified multiple ligand-binding sites for hGPR40 (ref. 6). Similarly Ca^{2+} flux assay also suggests a synergistic effect between TAK-875 (partial agonist) and γ -linoleic acid (full-agonist) binding and shows positive cooperativity between the two ligands³. In addition, point mutations of Arg-183^{5,39} Ala, Arg-258^{7,35} Ala, Tyr-91^{3,37} Ala and Asn-244^{6,55} Ala also show differential effects on the potencies of TAK-875 and γ -linoleic acid³. These observations strongly indicate that hGPR40 possesses more than one ligand-binding site.

Visual inspection of the receptor surface suggests two extra binding pockets in the receptor in addition to the binding site of TAK-875 (Fig. 3). One of the proposed additional binding pockets (site 2) is adjacent to TAK-875 and would potentially allow a ligand to pass between transmembrane helix 4 and transmembrane helix 5 (Fig. 3). Transmembrane helix 4 in hGPR40 is missing the frequently conserved proline^{4,60}, and concomitantly near the equivalent position there is a glycine residue, Gly 139^{4,58}. This may impart some degree of flexibility to the end of transmembrane helix 4, especially in the absence of TAK-875. Site 2 may also support entry of an additional ligand through the lipid bilayer. A potential third binding pocket (Fig. 3, site 3) is located near the traditional GPCR orthosteric site between transmembrane helix 1 and transmembrane helix 7. Interestingly, this location is very similar to where the positive allosteric modulator LY2119620 binds to the M2 receptor²⁴. Site 3 does not have a clear route through the membrane, which suggests that if this potential site is populated, it would bind ligands from the extracellular surface akin to earlier receptor structures. From the hGPR40 structure

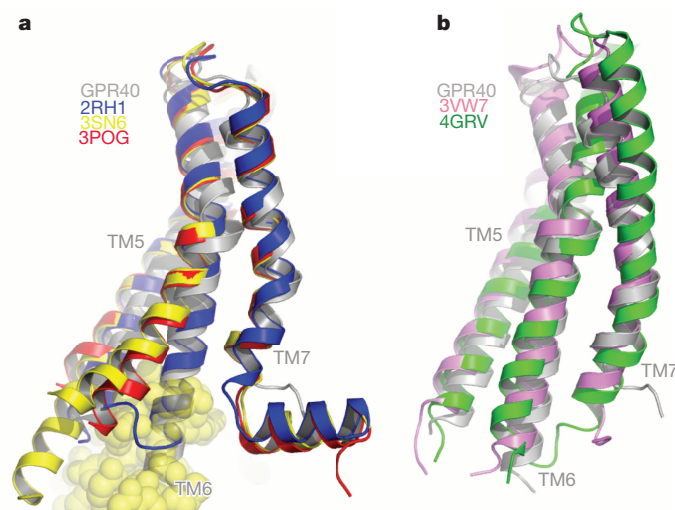


Figure 4 | Structure alignment of active and inactive GPCR structures. Structures are aligned based on equivalent residues in transmembrane helices 1–4. **a**, On the basis of the compendium of activation states of β_2 AR (active G-protein-bound 3SN6 in yellow cartoon with yellow spheres for the bound Gs, active-like nanobody bound 3POG in red, and inactive antagonist bound 2RH1 in blue), hGPR40 (grey) transmembrane 5 and transmembrane 6 adopt an orientation most like the inactive (antagonist bound) state. **b**, A comparison with structurally similar peptide-binding receptors, NTSR1 (active-like 4GRV in green) and PAR1 (inactive 3VW7 in magenta), suggests receptor subclasses may have a significant impact on transmembrane 5 and transmembrane 6 orientation regardless of activation state.

and available biochemical data, it is unclear whether all the three ligand-binding sites can be occupied at the same time or if they bind separately, each binding event further amplifying the agonist signal. Additionally, it is also intriguing if the receptor has multiple distinct fatty-acid binding sites. Current biochemical data and hGPR40 structure provide the groundwork for further studies to clarify the binding modes of fatty acids of varying chain lengths and other agonists.

GPCRs function in a dynamic equilibrium between multiple conformational states including inactive (R , R'), active (R'' , R^*) and signalling (R^*G) states^{25,26}. A comparison of the structure of the transmembrane helical region of hGPR40 with other agonist^{11,27,28} and antagonist¹⁵ bound GPCRs (Fig. 4) supports the notion that the hGPR40 receptor is in an agonist-bound R' inactive-like state^{25,26}. Unlike other active state agonist-bound GPCRs, transmembrane helix 6 in the hGPR40 structure does not reposition away from the helical bundle core. However, such as in the recently determined structure of PAR1, the C-terminal portion of transmembrane helix 7 adopts a conformation more consistent with the agonist bound β -2 adrenergic active structures^{12,27–30} (Fig. 4). Interestingly, this position of the C-terminal region of transmembrane helix 7 is observed in some of the other peptide-binding GPCRs. Future investigations may reveal if this feature reflects the active state of GPCRs or is a characteristic of peptide/lipid binding receptors.

TAK-875, an ago-allosteric modulator of hGPR40, represents a class of partial agonists that target receptor-mediated insulin secretion in a glucose-dependent manner. The structural information of the drug-receptor complex highlights the unique features of the TAK-875 binding pocket, including the charge network responsible for binding the carboxylate of the partial agonist. The structure also potentially provides insight into lipid entry of the ligand, and structural conservation between peptide- and lipid-binding receptors. Additionally, our structure presents a model of how multiple ligands could bind to the receptor to amplify the agonist signal.

METHODS SUMMARY

hGPR40 receptor with the mutations Leu-42^{2,40} Ala, Phe-88^{3,34} Ala, Gly-103^{3,49} Ala and Tyr-202^{5,58} Phe, and the ICL3 replaced with T4L, was expressed in *Spodoptera frugiperda* (Sf9) insect cells. The receptor was extracted in dodecyl- β -D-maltoside and purified by two-step affinity chromatography in a neopentyl glycol detergent for crystallization in lipidic cubic phase (Extended Data Fig. 1a–c). X-ray diffraction data were collected at beamline 5.0.3 of the Advanced Light Source at the Berkeley Center for Structural Biology, Lawrence Berkeley National Laboratory and at beamline 23-ID-D of the Advanced Photon Source at GM/CA, Argonne National Laboratory. The structure was solved by molecular replacement, and refinement statistics and the Ramachandran plot outliers as calculated by MolProbity are presented in Extended Data Table 2.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 October 2013; accepted 19 May 2014.

Published online 20 July 2014.

1. Itoh, Y. *et al.* Free fatty acids regulate insulin secretion from pancreatic β cells through GPR40. *Nature* **422**, 173–176 (2003).
2. McGarry, J. D. & Dobbins, R. L. Fatty acids, lipotoxicity and insulin secretion. *Diabetologia* **42**, 128–138 (1999).
3. Yabuki, C. *et al.* A novel antidiabetic drug, fasiglifam/TAK-875, acts as an ago-allosteric modulator of FFAR1. *PLoS ONE* **8**, e76280 (2013).
4. Negoro, N. *et al.* Discovery of TAK-875: a potent, selective, and orally bioavailable GPR40 agonist. *ACS Med. Chem. Lett.* **1**, 290–294 (2010).
5. Burant, C. F. *et al.* TAK-875 versus placebo or glimepiride in type 2 diabetes mellitus: a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet* **379**, 1403–1411 (2012).
6. Lin, D. C.-H. *et al.* Identification and pharmacological characterization of multiple allosteric binding sites on the free fatty acid 1 receptor. *Mol. Pharmacol.* **82**, 843–859 (2012).
7. Flodgren, E. *et al.* GPR40 is expressed in glucagon producing cells and affects glucagon secretion. *Biochem. Biophys. Res. Commun.* **354**, 240–245 (2007).
8. Kotarsky, K. *et al.* A human cell surface receptor activated by free fatty acids and thiazolidinedione drugs. *Biochem. Biophys. Res. Commun.* **301**, 406–410 (2003).

9. Shapiro, H. *et al.* Role of GPR40 in fatty acid action on the β cell line INS-1E. *Biochem. Biophys. Res. Commun.* **335**, 97–104 (2005).
10. Hirozane, Y. *et al.* Generating thermostabilized agonist bound GPR40/FFAR1 using virus-like particles and label-free binding assay. *Mol. Mem. Biol.* (in the press).
11. White, J. F. *et al.* Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).
12. Lebon, G. *et al.* Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **474**, 521–525 (2011).
13. Hanson, M. A. *et al.* Crystal structure of a lipid G protein-coupled receptor. *Science* **335**, 851–855 (2012).
14. Granier, S. *et al.* Structure of the δ -opioid receptor bound to naltrindole. *Nature* **485**, 400–404 (2012).
15. Zhang, C. *et al.* High-resolution crystal structure of human protease-activated receptor 1. *Nature* **492**, 387–392 (2012).
16. Wu, B. *et al.* Structures of the CXCR4 Chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
17. Wu, H. *et al.* Structure of the human κ -opioid receptor in complex with JDTic. *Nature* **485**, 327–332 (2012).
18. Hildebrand, P. W. *et al.* A ligand channel through the G protein coupled receptor opsin. *PLoS ONE* **4**, e4382 (2009).
19. Mercier, R. W. *et al.* hCB2 ligand-interaction landscape: cysteine residues critical to biarylpyrazole antagonist binding motif and receptor modulation. *Chem. Biol.* **17**, 1132–1142 (2010).
20. Hurst, D. P. *et al.* A lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. *J. Biol. Chem.* **285**, 17954–17964 (2010).
21. Negoro, N. *et al.* Optimization of (2,3-dihydro-1-benzofuran-3-yl)acetic acids: discovery of a non-free fatty acid-like, highly bioavailable G protein-coupled receptor 40/free fatty acid receptor 1 agonist as a glucose-dependent insulinotropic agent. *J. Med. Chem.* **55**, 3960–3974 (2012).
22. Sum, C. S. *et al.* Identification of residues important for agonist recognition and activation in GPR40. *J. Biol. Chem.* **282**, 29248–29255 (2007).
23. Sum, C. S. *et al.* Two arginine-glutamate ionic locks near the extracellular surface of FFAR1 gate receptor activation. *J. Biol. Chem.* **284**, 3529–3536 (2009).
24. Kruse, A. C. *et al.* Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101–106 (2013).
25. De Lean, A. *et al.* A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β -adrenergic receptor. *J. Biol. Chem.* **255**, 7108–7117 (1980).
26. Katritch, V. *et al.* Structure-function of the G protein-coupled receptor superfamily. *Annu. Rev. Pharmacol. Toxicol.* **53**, 531–556 (2013).
27. Rasmussen, S. G. F. *et al.* Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–555 (2011).
28. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **469**, 175–180 (2011).
29. Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.* **372**, 1179–1188 (2007).
30. Choe, H. W. *et al.* Transmembrane signaling by GPCRs: insight from rhodopsin and opsin structures. *Neuropharmacology* **60**, 52–57 (2011).

Acknowledgements We thank K. P. Wilson, M. Hixon and K. Goodwill for review and feedback on the manuscript; T. Ho and S. Okubo for molecular cloning support; T. Sjöberg for membrane preparation; and C. Dillard for cell culture support. We also thank the staff of the Berkeley Center for Structural Biology, Lawrence Berkeley National Laboratory, which operates Advanced Light Source beamline 5.0.3, and the staff of GM/CA, Argonne National Laboratory, which operates Advanced Photon Source beamline 23ID-D. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health and National Institute of General Medical Sciences. The GM/CA has been funded in whole or in part with Federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-1104). The Advanced Light Source and Advanced Photon Source are supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under contract numbers DE-AC02-05CH11231 and DE-AC02-06CH11357, respectively.

Author Contributions A.S. guided construct and expression optimization, developed purification procedures and purified receptor protein, characterized crystallization constructs, performed crystallization and structure analysis, and was responsible for project strategy. J.Y. performed crystallization and crystal harvesting, structure determination and structure analysis. Y.H. performed mutant screening and characterized mutants and crystallization constructs. G.S. and W.L. performed data collection, data processing and structure analysis. G.K. performed molecular biology, construct optimization and expression screening. F.G. performed construct characterization, protein purification and crystallization. K.A. guided construct optimization and performed molecular biology. A.J. and A.I. performed homology modelling to aid molecular refinement. A.J. discussed ligand-binding site in the structure. J.N. performed protein purification and crystallization; K.O. supported experimental design and data discussion for mutant screening. F.G., A.S., J.Y., G.K., G.S. and W.L. prepared the figures. All the authors contributed to manuscript writing.

Author Information Coordinates and structure factors have been deposited in Protein Data Bank under accession number 4PHU. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.S. (anika.srivastava@takeda.com).

CORRIGENDUM

doi:10.1038/nature13610

Corrigendum: Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET

Toshiyuki Matsui, Danny Leung, Hiroki Miyashita, Irina A. Maksakova, Hitoshi Miyachi, Hiroshi Kimura, Makoto Tachibana, Matthew C. Lorincz & Yoichi Shinkai

Nature 464, 927–931 (2010); doi:10.1038/nature08858

In Fig. 1a of this Letter, it has come to our attention that the lanes for the *Dnmt3l*^{+/+} and *Dnmt3l*^{-/-} testes RNA samples were run on different gels (northern blots) but were not displayed as such. The *Dnmt3l*^{-/-} sample was originally intended as a positive control for derepression of different retroelements analysed in this paper, such as LINE-1 and IAP. However, owing to space limitations, these samples were only mentioned in the legend to Fig. 1a. Regardless, according to Nature’s publication guidelines, these lanes should have been displayed separately. In addition, we have now inserted a black line between lanes 4 and 5 (counting from the right-hand side) of Fig. 2a and lanes 1 and 2 (counting from the right-hand side) of Supplementary Fig. 3a because although they were run on the same gel, they were not adjacent in the original gel. The Supplementary Information of this Corrigendum shows the corrected Supplementary Fig. 3a and the raw data used to prepare Figs 1a and 2a and Supplementary Fig. 3a. These corrections do not change our interpretation of the results or the conclusions drawn in this Letter. Nevertheless, we apologize for any confusion the original display may have caused.

Supplementary Information is available in the online version of this Corrigendum.

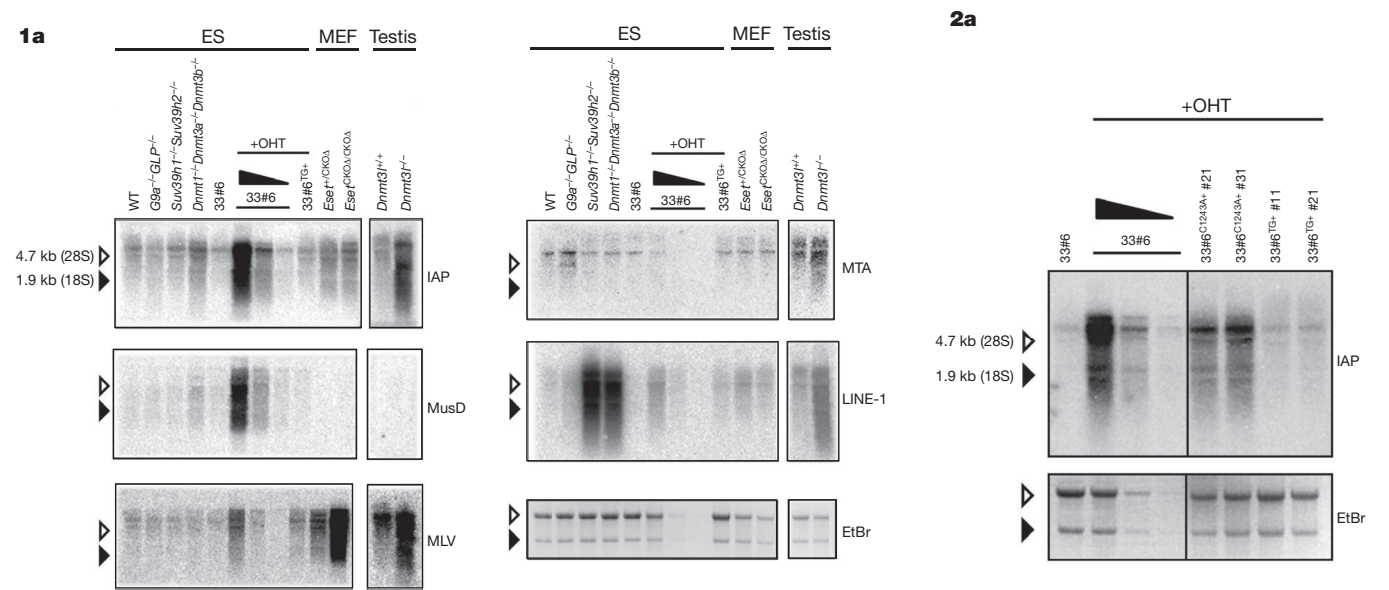


Figure 1 | This figure shows the corrected Figs 1a and 2a of the original Letter.

TOOLBOX

HOW TO TAME THE FLOOD OF LITERATURE

Recommendation services claim to help researchers keep up with the most important papers without becoming overwhelmed.

ILLUSTRATION BY THE PROJECT TWINS



BY ELIZABETH GIBNEY

Casey Bergman's daily research routine used to include checking all his e-mails and web alerts to pick out fresh papers in his field. But he grew dissatisfied with table-of-contents alerts from journals, RSS (Rich Site Summary) feeds and automated e-mails from the PubMed database. The flow of content was manageable, but if he left it for more than a day, "it became a burden", he says.

So last year Bergman, a computational geneticist studying fruit flies (*Drosophila*) at the University of Manchester, UK, turned to a fresh approach: an automated Twitter account (or 'twitterbot') that he named FlyPapers. The bot trawls PubMed and the arXiv preprint server to find papers containing the word *Drosophila*, and spits them out into its followers' feeds. Bergman finds it much easier to catch up with FlyPapers popping up in his Twitter feed — and his idea has spawned

around 55 twitterbots in other disciplines.

It is no surprise that academics are coming up with their own ways to keep on top of the flood of literature. "It's a common struggle," says Bergman. A staggering 6,000 papers are published every day — and although no one wants to be overloaded with recommendations, missing key papers is "mortifying", says Sally Burn, a developmental geneticist at Columbia University in New York City. She uses a service called Scizzle, which regularly sends her the results of saved

RECOMMENDATION ENGINES

*A guide to reading***Google Scholar** (scholar.google.com)

Sends alerts about recommended papers on the basis of a user's publication history.

ReadCube (readcube.com)**Mendeley** (mendeley.com)

Reference managers with recommendation engines (Readcube is backed by Digital Science, a sister firm to Nature Publishing Group).

PubChase (pubchase.com)

Recommends papers based on the libraries of users with similar interests.

Sparrho (sparrho.com)

Asks the user to train its recommendations engine by approving or rejecting suggestions.

Faculty of 1000 Prime (f1000.com/prime)

Sends alerts about biomedical articles, using the ratings of 5,000 senior scientists.

Twitter (twitter.com)

Automated twitterbots can track keywords (see go.nature.com/bppzew for instructions), or users can follow colleagues, as they can on other social networks such as ResearchGate.

Nowomics (nowomics.com)

Users 'follow' biological keywords such as specific genes, proteins or processes.

Scizzle (myscizzle.com)

Automates the process of making multiple PubMed searches with keywords and filters, and allows users to bookmark relevant papers.

PubMed searches. "Unless you have all day, and ten people working for you trawling the literature, I think it's the best situation you're going to get," she says.

But a stream of papers based on keywords only scratches the surface of what is technologically possible. Emerging literature-recommendation engines promise not only to filter the flood of papers to a trickle, but also to learn from their users' interests to add personalized suggestions (see 'A guide to reading'). "In spirit, it's similar to what Netflix or Amazon do," says Matthew Davis, a computational biologist at the University of Texas at Austin who wrote the algorithm for one such service, PubChase — now owned by ZappyLab, a firm in Berkeley, California, that makes web- and phone-based tools for scientists.

IF YOU LIKE THAT, YOU'LL LIKE THIS

One of the first, and still best-known, services comes from Google Scholar. Its Updates tool suggests articles by applying a statistical model to a record of a researcher's authored papers and citations. "The recommendations are almost scarily good," says Roger Schonfeld, programme director at Ithaka S+R, a non-profit consultancy based in New York City that advises academia on digital technology. But graduate students may not have a sufficient body of work for the site to help, notes Patrick Mineault, a computational neuroscientist at the University of California, Los Angeles.

PubChase suggests articles from PubMed on the basis of a user's publishing record, but it also learns from the articles that the user has read and stored in his or her online library. And it adds another machine-learning technique: comparing this library with other people's collections, with the logic that people with common research interests might benefit

from each others' preferences. "I've been really impressed: nearly every article it has recommended has been relevant to my research," says Kelsey Wood, a geneticist at the University of California, Davis, who uses the service along with reference-manager tool Mendeley, owned by Amsterdam-based publisher Elsevier.

Ross Mounce, an evolutionary biologist at the University of Bath, UK, says that PubChase is not useful for those whose interests fall outside the boundaries of PubMed. He prefers Sparrho, a fledgling London-based venture that generates recommendations with a keyword-based feed, and asks users to train the tool by flagging suggestions as relevant or irrelevant. It

includes articles, grants, patents, posters and conference proceedings from all the sciences. "The breadth is a real strength," says Mounce. As with PubChase, recommendations are based on connections between similar users. "We're allowing intelligent curators, humans, to join the scattered dots," says chief executive Vivian Chan, who co-founded Sparrho after she struggled to keep up with the literature while studying for a biochemistry PhD at the University of Cambridge, UK.

As start-ups seeking investment, PubChase and Sparrho are guarded about how many users they have. It is clear that numbers are small. (A *Nature* survey of more than 3,000 scientists found that only 8% had heard of PubChase, and fewer than 1% visited it regularly; see *Nature* **512**, 126–129 (2014).) But both say that their user base is growing quickly.

BACK TO BASICS

Bergman is wary of algorithm-based searches. A machine that learns and tailors recommendations can become like "blinders on your intellectual scope," he says. And he has found that the interdisciplinary nature of his work, which melds genomics and text-mining, confused Google Scholar — the tool threw up irrelevant papers and missed important ones. But Davis says that this narrowing is counteracted by the new doors opened by recommendations based on the profiles of people with similar interests.

Many researchers eschew algorithms altogether, and simply follow colleagues on social networks to find out what is worth reading. "Twitter is the unsung hero of the paper-recommendation world," says Cassie Ettinger, a geneticist in the same research group as Wood. Other scientists check which papers rise to the top in online communities or among users of reference-management services such as Faculty of 1000 Prime and Mendeley.

But the desire to share recommendations or upload libraries to find new papers is hardly universal. Derek Lowe, a chemist at Vertex Pharmaceuticals in Boston, Massachusetts, who writes the blog *In the Pipeline*, remains a fan of RSS feeds from journal websites. And Burn says that she does not have the time to train a recommendation engine. Mineault acknowledges that automated learning devices will never find all the papers a scientist wants, but he thinks that they will improve. Techniques for gleaning meaning from content will become more sophisticated, he says, and will eventually have a significant role in guiding scientists' reading choices.

For Bergman, a lot of this is a matter of taste. His twitterbot has convened an online fruit-fly community; its suggestions have been retweeted by researchers in other disciplines, and even by non-scientists. Bergman has not ruled out trying further technologies, but he is sticking to FlyPapers for now. "I haven't felt the need to try any others. It's working for me, and that's all that matters," he says. ■ **SEE EDITORIAL P.6**

MORE ONLINE**Q & A**

SHUTTERSTOCK/ISAK55

Software carpentry: the boot camp for scientific software. go.nature.com/eijvxi

ARTICLE COLLECTIONS

- Culture of code: are scientists' practices up to the task?
- Online collaboration: what makes a winning formula?
- Open data: how to share research online.

For more on scientific software, apps and online tools, visit nature.com/toolbox

CAREERS

TURNING POINT Brown-fat work wins kudos for obesity researcher **p.133**

NATUREJOBS BLOG Discussions and tips on jobs and careers. blogs.nature.com/naturejobs

NATUREJOBS For the latest career listings and advice www.naturejobs.com

RETROCKET/VETTA/GETTY



INTERVIEWS

The all-important job talk

How people present their research and aspirations during interviews for academic positions can make or break their career.

BY AMANDA MASCARELLI

When Spencer Smith was being interviewed for a faculty post in neuroscience four years ago, he told his audience a story about the British race-car driver Stirling Moss.

In 1955, while Moss was training for the Mille Miglia, a famous thousand-mile endurance race in Italy, he realized that when he was intensely focused on driving, he could not hear the instructions given by his navigator, Denis Jenkinson. Jenkinson therefore

used hand signals to guide Moss, and the pair won the race, setting a record time that still stands today.

Smith had used the story about Moss to illustrate his research interests on how various parts of the brain work together to support behaviour. Thanks at least in part to that lively story, he landed the position, at the University of North Carolina in Chapel Hill — and his presentation is still well remembered. “You remember his science because you remember that story,” says Kathleen Caron, who chairs the department

of cell biology and physiology.

At many universities and institutes worldwide, candidates for jobs in academia go through a two-day interview process that includes a formal, public, 45-minute talk — informally known as the ‘job talk’ — followed by a question-and-answer session. It usually takes place in an auditorium or lecture theatre that seats 50 to several hundred people, and the audience typically includes not only members of the hiring committee, but also graduate students, postdocs and faculty members from the hiring department and ►

► often other departments too. The talk may be just one piece of a candidate's application, but it carries enormous weight. Those who have succeeded — and who have watched candidates give talks that shine or that fall flat — say that there are ways to avoid common mistakes and to leave a positive lasting impression (see 'Tips for success').

Talks at conferences, seminars and departmental meetings generally aim to convey research findings and foster questions, but a job talk must do more. It must give a sense of amicability, convey enthusiasm, illustrate the research vision and show why the candidate would be an excellent fit in the department.

A poorly rendered talk can mean no job, says Stefano Stifani, a neuroscientist at the Montreal Neurological Institute and Hospital in Quebec, Canada. "It puts the thoughts in people's minds: 'Will this person really be a good teacher? And does this person really get the big picture?' Whether that is the case for a candidate is a key factor in predicting whether the person will write good grants."

Caron agrees. "It's the most important 45 minutes of your career," she says. "You sink or swim on that job talk." Candidates will be judged on "the whole package" they bring, she explains. "You're showing how you're going to make a scientific argument. What are you going to do in the future? Do you have the skill set to convince people that you should be funded? And do you do that in a charismatic, engaging, intellectually stimulating and fun way that makes you a nice colleague?"

PEOPLE PERSUASION

A central mission of the talk is to demonstrate compatibility with potential colleagues. James White, a palaeoclimatologist at the University of Colorado, Boulder, says that he does not

understand why department members do not flock to these talks. "I always like to have some knowledge of the person who I'm going to be sitting with for the next 30 years," he says. "Is there any rapport? Is this somebody I can get along with? Is this somebody who's going to be a pain?"

Stories such as Smith's that resonate with the audience and reveal an engaging personality can be very effective in winning over a hiring committee.

"Too often in professional talks, we're afraid to let who we really are show through," says Alan Townsend, who in July became dean of the Nicholas School of the Environment at Duke University in North Carolina. "Some people are naturally very funny, so be funny. But if you're not, it doesn't mean that you can't still connect in really powerful and meaningful ways on a very human level."

During his job talk, Townsend shared with his audience that he had been a finalist for the position seven years before. This time around, he joked, he was going to "shake things up and try something new".

He told a humorous story about two classmates who had competed for class president when he was in school. The student who won the election had sung a few lines of a tune by the classic US rock band Cheap Trick, *I Want You to Want Me*. Townsend gave his own rendition, and in doing so, earned laughs and applause — along with the job.

The right level of detail is also key. "One of the most common mistakes that people make is that they fail to engage the broad

audience," says Stifani. "Present your data in a way that is accessible, even to non-specialists." The audience for a typical job talk at Stifani's institute might include faculty members from sub-disciplines such as neurodevelopment, neuroimaging or neurodegeneration. But it might also include a cancer biologist or an immunologist from another institute. "Everybody has to be able to extract what the key significance was," he says. "If that doesn't come out, I think the candidate is toast."

But there is a tricky balancing act to be achieved — candidates also need to capture the sophistication needed to satisfy specialists. If a talk is accessible to non-specialists, experts in the field might complain that it was too shallow, and vice versa, says White. "I have seen some talks that were scientifically brilliant," he says, but some hiring-committee members wondered whether the candidate would be able to give an equally effective talk to non-experts. To address this common dilemma, White's department now often asks candidates to also give a 'teaching talk', in a smaller, more informal setting, to allow the hiring committee to assess the candidate's teaching and outreach skills.

Candidates should find out in advance what the hiring department is expecting by talking to the chair of the search committee, the department chair and sometimes to other faculty members.

Early in the job talk, candidates should state their key research questions and why those matter. "I like to see them choose important problems, and I like to see them articulate that up front," says White. The heart of the talk — the science — is "the thing that's hardest to fake", says Jonathan Payne, a palaeontologist at Stanford University in

TIPS FOR SUCCESS

How to avoid common pitfalls

DO:

- Practise, practise, practise. Find people, whether through your home department or through social media, who would be interested in hearing your practice talks and offering feedback.
- Attend plenty of job talks and departmental seminars and pay attention to what works well and what does not.
- Watch and study talks sponsored by the non-profit organization TED (short for technology, entertainment and design). They are up to 18 minutes long and have become widely recognized as powerful platforms for issues of universal appeal.
- Look at the website helpmeviz.com for in-depth feedback on data visualization.
- Use a smartphone, iPad or video camera

to record yourself practising and save the files. Once you find a version you are satisfied with, listen to it over and over again.

- Explain clearly what your research questions are, why you are doing the work and why it is important. Tell the audience what you are going to tell them, then tell them, then recap what you told them.
- Rehearse several ways to end your talk, in case time is running out and you have to wrap things up quickly.
- Familiarize yourself with the room and with audio-visual and other equipment in advance.
- Use multimedia and special effects judiciously.
- Play to your strengths and be yourself. Do not force humour. If you use it, keep it

light. Clear, concise, authoritative talks are compelling in their own right.

- Demonstrate that you know what other people in the department are doing, and how your research would fit in. That shows you have done your homework, and it gives hiring committees a clear idea of where collaborations might develop.

DO NOT:

- Do a 'data dump'. Resist the urge to show a lot of data — you lose audience members by overwhelming them with minutiae.
- Exceed your time allowance. By doing so, you signal that you are underprepared, did not take the presentation seriously or are arrogant. Try to end early so that there is sufficient time for questions. **A.M.**

California. “You need to be working on a good problem and you need to have some interesting results,” he says.

Candidates should also give a clear road map of where they are headed with their research, says Eileen Furlong, head of the Genome Biology Unit at the European Molecular Biology Laboratory in Heidelberg, Germany. She recommends that job candidates reserve the last 20–25% of their talk for discussing research plans and showing why she or he is the best person to address those questions.

A sign of success, says Furlong, is when audience members “leave the room thinking, ‘That’s really interesting and exciting, and I want to talk to that person.’”

DEPARTMENT DISCUSSION

The talk should also convey how well the candidate would fit into the department, says Amanda Clark, a neuropsychologist at the University of Tennessee in Chattanooga who interviewed for her faculty position in 2012.

By looking at the university’s website and promotional materials as well as other faculty members’ programmes, Clark knew that the university valued people who would offer undergraduates some research experience. She therefore wove in details about how undergraduates and graduate students had participated in her research. She also knew that the department was looking for someone with expertise in biological psychology and neuroscience but did not have an animal lab or an imaging facility. So in discussing her research on brain trauma and neurodegeneration, she highlighted her use of human subjects in her studies.

Candidates also need to prepare for the unexpected, such as technical glitches, a pushy audience member during the question-and-answer session and even a power outage.

Caron was halfway through her talk for tenure — another key time where talks can make or break a candidate — and was about to tell the audience about her most exciting and crucial piece of data. Suddenly, the electricity went out. Worse yet, the talk was in a room with no windows, so the crowd of 75–100 people was left in darkness. “It was absolutely pitch black,” says Caron. She had no choice but to wrap up her talk quickly.

“Then the most incredible thing happened,” she says. “People asked questions. I was so touched by that. We all sat in the dark room and talked about science, which was really wonderful.”

Even better, Caron got tenure. ■

Amanda Mascarelli is a freelance journalist in Denver, Colorado.

TURNING POINT Shingo Kajimura

Diabetes researcher Shingo Kajimura studies how fat cells can be manipulated to control obesity. In March, he was fêted for winning one of the US Presidential Early Career Awards for Scientists and Engineers (PECASE). He explains how favourable timing has helped him to advance his career.

Did you always plan to become a researcher?

No. Growing up in the suburbs of Tokyo, I went fishing every day and loved to see living things. I was planning to try to work with animals — as a national park ranger at a zoo or an aquarium, for instance. No one in my family had ever attended university or earned a degree, so I had not even known that the job of scientist existed until I started there. But I had good mentors at the University of Tokyo, who guided me towards science. I was surprised by how much I loved it.

What led you from Japan to the United States?

I won fellowships that allowed me to spend my entire graduate time in the United States, and spent two years in Hawaii and three years in Michigan studying fish physiology and developmental biology. I thought about going back to Japan after that, but I felt like I could be more intellectually independent in the United States, and that was important to me. Then a Pathway to Independence award from the US National Institutes of Health (NIH) gave me the confidence to become a principal investigator here.

What led you to study brown fat?

I received a fellowship at Harvard Medical School in Boston, Massachusetts, to examine how mitochondria help living things to adapt to different environments and whether that process could be manipulated. When brown fat was discovered in adults, it opened up new opportunities for research into obesity and related diseases such as insulin-resistant diabetes. I decided to focus on brown fat in mammals because it is one of the most mitochondria-rich cell types in the body, which means it has a big role in regulating the body’s adaptation to different environments.

How did the discovery of brown fat in adults affect your own career?

When I started the postdoc in 2006, the brown-fat research field was pretty inactive because nobody knew adults had brown fat. When that was discovered in 2009, it led to additional funding for brown-fat research. I was trying to see whether we could convert white fat cells into brown fat genetically or chemically. In 2009, I had a paper published in *Nature* about



the body’s ability to convert some cell types to brown-fat cells (S. Kajimura *et al. Nature* **460**, 1154–1158; 2009), a process now known as direct reprogramming, as well as papers published in other journals. The following year, I received the NIH grant and landed a post as an assistant professor at the University of California, San Francisco.

What does the PECASE award mean for you?

The money helps, and the award has also connected me with many talented young scientists in different fields for potential collaborations.

What are your biggest challenges right now?

Even if I earn tenure, I still have to get grants, and that is tough. And I have gone from being a player in the lab to a coach. Nobody gets training in how to manage lab staff, and everyone does it in their own way. For graduate students with less experience in bench work, I start them with relatively easy experiments; for postdocs, I encourage them to help on projects close to the finish line so that they can co-author a paper. It helps them get used to the new environment and to work as a team member.

What is the best advice you have ever received?

One of my mentors always said that if the science goes well, everything else will follow. So I try to spend as much energy as I can on the science and not worry about other things. It is not always easy, but I remind myself that it is really about the science.

What is the secret to your scientific success?

Internal passion. I am not a genius but I genuinely love science, so I don’t feel like this is my job really. ■

INTERVIEW BY SCOTT KRAFT

INVESTMENTS

Creative thought.

BY SIMON KEWIN

Evangelina Carter, chief executive of Blue Planet Holdings, stared at her visitor from Head Office. She must have misheard his words. "I'm sorry. For a moment I thought you said you wanted to wipe out civilization."

Mr Allen peered over his half-moon glasses. You had to give it to them. They had their human mannerisms down to a T. "No of course not. Not *wipe out*. We merely wish to ... subdue humanity. Knock it back to a less technological era. Our projections suggest five centuries should do the trick."

Evangelina didn't speak for a moment. They could do it, too. She studied the ancient alien sitting opposite her. In his finely tailored suit and old-school tie he looked every inch the genial investment banker. Allen smiled, awaiting her reply. Beyond him, through floor-to-ceiling windows, London stretched away into a hazy distance.

With an effort she maintained her professional manner. There was a place for panic and terror but the boardroom wasn't it. "But, why? Is there a problem with output?"

"Heavens, no. We are most satisfied with your stewardship, Ms Carter. Profits from operations on Earth continue to soar. The creativity of this planet remains as phenomenal as ever. But we've made projections and the results are clear. Another century or two and things will be different. We have to protect our investments."

She dreaded asking. Their forecasts were always accurate. "May I know what you foresee? Environmental collapse? Wars? Pandemic?"

"Quite the contrary. Humanity will thrive. Scientific advance will continue apace, ushering in a new golden age of abundance."

"Then I don't see the problem."

Allen looked troubled. It was all for show. His face, like the rest of him, synthesized. So far as she could tell, her masters were amorphous blobs of jelly. But amorphous blobs of jelly with vast technological resources at their fingertips. If that was the right word.

"Ms Carter," he said, picking his way through his words, "has anyone ever explained why we chose your planet?"

"I assumed you monitored us for millions of years."

His face twisted into new heights of trouble. "Quite so," he said. "But the truth is we didn't simply monitor."

"You didn't?"

"No, we also ... shepherded. Guided. Intervened."

"Which is against galactic statute?"

"Yes. But our projections, you see. They were quite clear. Given the right conditions your remarkable species would produce — well, all the glories it has produced. I don't need to list them. Music, literature, film. The Galaxy can't get enough of it and we, as rights owners, make a fortune."

"So what did you do?"

"Ms Carter, I'm telling you this because we trust you, yes?"

That was something. "Go on," she said.

It was his turn to stare through the window. More play-acting. "Ms Carter, I hope you won't be angry when I inform you that every other sentient species in the Galaxy is, to all intents and purposes, well, *immortal*."

"Immortal."

"Quite so. We don't die. I myself have been alive for nearly 600,000 years. And I'm considered rather a young gun. A bit of a hot-head."

"Immortal," she said again, as if the word was unfamiliar. "What does that have to do with us?"

Allen switched to *deeply sad* now. "Well, you see, here's the thing. Immortality is lovely, of course, but it can be so ... enervating."

"Can it? How awful."

"Yes. Oh, we set out to achieve great things. Works of art and feats of literature. Musical masterpieces. But knowing you can set it aside for a thousand years — well, frankly, it's hard to motivate oneself."

"I can see that."

"But humanity, now. You blaze briefly but gloriously across the face of the Universe. Knowing you have only a few years focuses your minds wonderfully. You yearn for the eternal without really knowing why."

"Are you saying you did this to us?"

Allen took off his glasses and polished them on his silk handkerchief. "I'm afraid so. A few mutations introduced millennia ago. Your cellular

structures decay when they really shouldn't. A few decades of life and ... poof! You're gone. I really am sorry."

She was beginning to see where this was going. "And your projections?"

"Well, it's this damned science, you see. Another century or two and you'll be unravelling all our work."

Humanity will defeat death, and creativity will plummet. It'll wreak havoc with our profits, Ms Carter. Havoc!"

"So you plan to return us to somewhere around — what — the Renaissance?"

"Ah, the Baroque," said Allen. "Such musical glories."

She considered. "But you'll still need intermediaries. You'll need me?"

"Quite so. Our trusted agent."

Someone who understands local custom. We shall, of course, ensure that you're spared the horror as it engulfs Earth. Your title will have to change, too. You could become an Empress, say. Would that suit?"

What they'd done was monstrous. An evil so vast she couldn't even think of a name for it. She couldn't let them get away with it.

Except ... she loved being in charge of Blue Planet. And Empress Evangelina had a ring to it. "These genetic changes. They're reversible?"

Allen regarded her over his spectacles. "Undoing them would be terribly difficult."

"But possible?"

"Theoretically. But counterproductive."

"I don't mean everyone. I mean *me*."

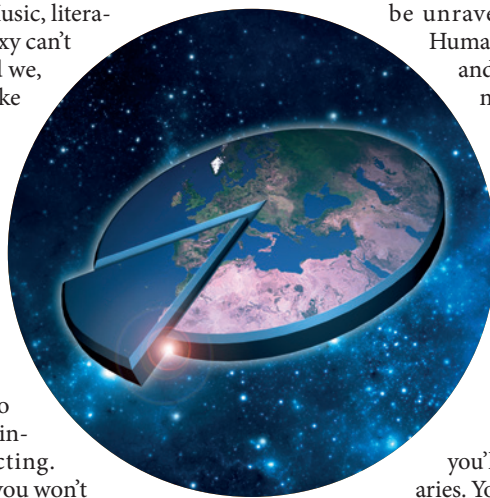
"You?"

She smiled. "It'll be our little secret. I can undergo the treatment while you're busy knocking humanity into the dark ages, yes?"

He hesitated for a moment, but she had him. For all their godlike technology they were useless at cutting deals.

After a moment he nodded and held out a synthesized hand. "Very well, Ms Carter. Let us shake. Our little secret." ■

Simon Kewin writes fantasy, SF, mainstream and some stories that can't make their minds up. His fantasy novel *Hedge Witch* has just been published. Find him at simonkewin.co.uk.



ON NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm